



Alexa Internet

Presidio of San Francisco
PO Box 29141
San Francisco, CA
94129-0141

tel: (415) 561-6900
fax: (415) 561-6795
frontdesk@alexa.com
www.alexa.com

ALEXA INTERNET PRESS CLIPS

June 1997 to April 1998



Digitized by the Internet Archive
in 2013

JULY

JULY

Main
Head
New

Alexa Launch Coverage: July, 1997

Headline:	Publication:	Date:	Author:
Alexa Keeps Surfers on Beaten Path	Inter@ctive Week (print)	7/28	Tom Steinert-Threlkeld
Alexa Searches While You Surf	PC Week	7/28	Jim Kerstetter
Revolution Brewing at Presidio: Alexa Internet hopes to change the way you use the Net	San Francisco Examiner	7/27	Zachary Cole
The Internet Boom Is Just Beginning: Big business is the next frontier, attendees at Chicago trade show say	San Francisco Chronicle	7/25	Jon Swartz
Outside The Box	San Jose Mercury News	7/24	Chris Nolan
As Go Surfers, So Goes Alexa	Wired News	7/24	Chris Oakes
2 Internet Giants Won't Show at Expo: Microsoft, Netscape are skipping Chicago	San Francisco Chronicle	7/23	Jon Swartz
New Ad-Backed Service to Monitor Context of Web Pages	Internet Advertising Report	7/23	Compiled by Beth Cox
New Guide Offers Net Directions	Inter@ctive Week (online)	7/22	Tom Steinert-Threlkeld
New Search Tool Tracks User's Path Through Internet	The San Diego Union-Tribune	7/22	New York Times News Service
NewsScan: Today's Headlines from the Mainstream Press	PC Magazine Online	7/22	NewsWatch

Headline:	Publication:	Date:	Author:
Looking for a Second Commercial Hit, Brewster Kahle Launches a Navigation Service	Web Week	7/21	Margaret McKegney
Alexa Internet Lands at the Presidio	San Francisco Business Times	7/21	Steve Ginsberg
New Service Tracks Web Use	Nando.net (Reprint of NY Times Article)	7/21	John Markoff
New Service Tracks Web Use	New York Times	7/21	John Markoff
Alexa Internet: The Search as a Communal Effort	New York Times (CyberTimes Extra)	7/19	Laurie J. Flynn
Alexa Browser Companion Makes Searches More Certain	PC World Online	7/18	Lisa Moskowitz
Re-engineering Surfing	San Jose Mercury News Online (link to TechWeb story)	7/18	Patricia Sullivan, Online Editor
Internet Historian Creates Tool to Inform Web Site Visitors	ZDNN-- The ZDNet News Channel	7/18	Renee Deger
Alexa Debuts Search Engine That Learns from Its Users	ZD Internet MegaSite Magazine	7/18	no byline
Startup Seeks To Synthesize Web	TechWeb	7/17	John Gartner
Alexa Internet Introduces Web Navigation that Learns from People	LA Times, BizWire Section	7/17	Press Release
Alexa Makes the Web History	C/Net	7/9	Alex Lash

BE ON THE CUTTING EDGE OF IT CONTRACT CAREERS

cnet

NEWS.COM

front page

Get now FREE!

MENU

[Front Door](#)

[The Net](#)

[Computing](#)

[Intranets](#)

[Business](#)

[CNET Radio](#)

[Perspectives](#)

[Newsmakers](#)

[Rumor Mill](#)

NEWS OPTIONS

[The Week View](#)

[Desktop News](#)

[News Alerts](#)

[Custom News](#)

[Advanced Search](#)

[Help](#)

click here.

**So you can't
figure out
the internet?**

advertisement

The Net

[back to](#)

FREE newsletter

[enter email](#)

Subscribe

Alexa makes the Web history

By Alex Lash

July 9, 1997, 1:05 p.m. PT

Confused or swamped Web surfers will soon have yet another tool at their disposal as the [Alexa Internet](#) Web navigation service goes into beta next week.

The new service is run by the creators of the [Internet Archive](#), a project that aims to document and store Web pages, Usenet records, and shareware for the historical record.

The Alexa client is a toolbar that sits at the bottom of a browser as the user surfs the Web. Once downloaded and installed, the Alexa client will use a [bot](#) to analyze each Web site visited. The information it gathers will then be displayed on the toolbar. Alexa will relate information on the site's traffic, where the site is physically located, the number of pages it contains, and when the site was last updated, according to a company spokesperson.

Alexa will also display relevant links to other sites by analyzing the paths of previous visitors. Alexa will add the URLs it analyzes to the Internet Archive.

The software will be available to download for free from the Alexa Web site. The company would not comment on how it plans to bring in revenue.

Latest Headline
[display on desktop](#)

The Net

• [Bells Labs finds J](#)

• [Alexa makes the V](#)

• [Netscape to use its](#)
[software](#)

• [Crown has taken a](#)
[step](#)

• [ETC's Winery to](#)
[practice](#)

Computing

• [Falling prices mean](#)
[memory](#)

• [Soft, some gleam i](#)

• [Wipro has highest](#)
[score](#)

• [New line of 200-A](#)

Intranets

• [Bay makes a high-](#)

• [Another deal comi](#)

• [ISPs may lose intr](#)

Business

• [Amelia Hancock](#)

• [Amelia Hancock](#)

• [Wayner's Wire ge](#)

• [Digital in talks to](#)
[business](#)

• [Trade group CEO](#)

• [Search engines urge](#)
[business move](#)

• [FCC goes Baby B](#)

RESOURCES

[service](#)



Friday
Jul 18, 1997

Tech Search

► [TechWire](#) ► [TechInvestor](#) ► [TechTools](#) ► [TechHelper](#)

TechWire

[\[Direct Article Link\]](#) [\[New Search\]](#) [\[Search the Web\]](#)

July 17, 1997, TechWire

Startup Seeks To Synthesize Web

By John Gartner

SAN FRANCISCO -- Startup **Alexa Internet** announced a product Thursday with lofty aspirations: to re-engineer the way people browse the Web. The Web-veteran founders of Alexa, located in the recently demilitarized Presidio compound in San Francisco, have unwrapped a browser companion application that focuses on the user experience instead of on text searching to guide people through the Internet.

The Alexa toolbar client provides users with detailed information about where they are and where they might want to go next. Alexa will distribute the client for free and will receive income by broadcasting ads based on the content currently being viewed.

For each site visited, the Alexa app displays the parent company and provides ratings for the depth and quality of information based on the amount of information and how long previous visitors remained at the site.

The "Where To Go" information suggests comparable sites based on where users most frequently branched off to next and provides a related advertisement. For example, if a user visits a Ford Mustang enthusiast page, Alexa might display a car advertisement and related sites on auto repair.

Alexa president and CEO Brewster Kahle is no stranger to grandiose designs for changing the way people access information. Kahle invented the Wide Area Information Server (WAIS) information retrieval system, which he later sold to America Online. He also worked as lead engineer on the Thinking Machines parallel-processing supercomputer project.

Privately funded, Alexa was founded in April 1996 by Kahle and fellow WAIS-veteran Bruce Gilliat.

Alexa's client/server architecture "sniffs" each user's trail around the Web and sends a copy of that path back to Alexa servers. This does not threaten individual's privacy, Kahle said, because only the pathway is recorded and not who's browsing.

This extra data requires less than 10 percent of a user's bandwidth, Kahle said. Once installed, the toolbar starts up with the browser, but it can be selectively turned off. The suggestions of where to go are based on Alexa's archive of the Web that is continually reindexed and re-evaluated. Processing the 2 terabytes of Web data is no big deal for someone used to working in the supercomputing arena, Kahle said. "The Internet is finally big enough to look at."

In addition to the surfing assistant, Alexa includes a real-time chat client for instant communication with other Alexa users.

Alexa stores and processes the data at its headquarters using racks of Sun SparcStations and

In a game where love means nothing,

Thursday, July 17, 1997

Alexa Internet Introduces Web Navigation That Learns From People

■ Free Toolbar Adds Context, Continuity and Efficiency to Web Searching

[PREV STORY](#)

NEXT STORY

SAN FRANCISCO--Alexa Internet announced today a new Web navigation service that learns from people. The service, Alexa, is free to users of the World Wide Web and provides a continuous source of relevant recommendations of where to go next on the Web. These recommendations are derived by analyzing patterns within the Web and anonymous usage paths. Alexa works in parallel with, and independently of, Web browsers and is displayed as a toolbar on the user's desktop. Users can download the beta version of the service starting today at <http://www.alexa.com/>. "Alexa learns from people's use of the Net, so that future users can find the good stuff they know is out there," said Alexa president Brewster Kahle. "In 1993, the Web allowed everyone to become a publisher; in 1997, Alexa will, in effect, allow everyone to become an editor."

Alexa: Know Where You Are and Where To Go Next

The free Alexa service is a toolbar that gives you information about where you are on the Web and where to go next. The client works with Web browsers and is displayed at the bottom of your screen. Alexa offers immediate, tangible benefits for Web users -- savvy and inexperienced alike.

For example, if you search the Web for mountain bike information you might first use your favorite search engine or directory. When you arrive at a mountain bike site, Alexa lets you know how others have liked this site -- whether they have backed out or decided to click through -- and what paths through the site most users followed.

If you decide to compare other brands of bikes, you won't have to return to your search engine or directory and begin a new search.

Instead, Alexa offers continual access to a list of sites frequented by others interested in the topic. Alexa's archive of the Web virtually eliminates the risk of finding a dead link when a page is no longer available on the Web.

Alexa users have easy access to four primary features contained in the Alexa toolbar:

Where am I?

Alexa provides you with live information on every site you visit such as:

-- Who the site is registered to -- How many pages are on the site -- How many others point to this site -- How frequently the site is updated -- The site's popularity among our users

Where should I go next?

Alexa uses the content of the Web and the insights of its users to help you decide where to go next. By analyzing patterns within the Web and anonymous usage paths, Alexa provides you with a continuous source of relevant sites that search engines would not suggest.

What happened to my page?

Alexa is the only service that provides you with quick, automatic access to an archive of the entire public content of the Web. Now, instead of your browser displaying a "404 --

Not found" message. Alexa will retrieve a copy of the page from our archive.

Who else is online?

Alexa lets you know which of your friends and colleagues are online so you can communicate with them in real-time.

"Scuba divers use the buddy system when navigating underwater; cybersurfers need similar help," said Jerry Michalski, Managing Editor of Release 1.0. "Because it combines access to publicly available Internet information with useful analysis of member activity and smart use of Internet resources, Alexa offers such help. The service should make using the Internet simpler and more compelling for many people."

The Collective Wisdom of the Web -- People

Internet search engines perform searches based on a few keywords. These search engines provide no context when returning suggested links to the user. The breadth of the information that they have gathered and indexed is their strength; difficulty of use and over-generality are their weaknesses.

In comparison, Alexa's ever-changing and context-sensitive lists of suggested sites to visit are created from the analysis of the combination of the overall traffic patterns of the Web, the ecology of links between sites, the content of the Web pages and anonymous usage paths. In essence, Alexa harnesses the collective wisdom of the Web by taking advantage of its most killer app -- the people who use it.

"Alexa provides context-sensitive advertising with our 'Where to go next' suggestions," said Bruce Gilliat, General Manager of Alexa Internet. "While users will receive advertisements when using the service, the ads will be relevant to what you are viewing and in many cases may act as additional site suggestions." The beta period will be limited to the first 10,000 users to register and download the client. The service will then be unavailable for download until it is available for general distribution toward the end of this summer. Alexa will initially be compatible with browsers that run on a Windows95 or Windows NT operating system.

Alexa Internet Partners

Alexa is working with several strategic partners. Encyclopaedia Britannica and Alexa have agreed to integrate a suite of content services to provide a comprehensive desktop reference tool. Alexa and TRUSTe, a global, non-profit initiative to establish trust and confidence in electronic communication, are working together to give users more information to make wise Internet decisions.

Additionally, Alexa Internet is working with Aptex to apply the Aptex Convectus(TM) software to the problem of context-sensitive suggestions.

About Alexa Internet

Alexa Internet analyzes multi-terabyte collections of data to create navigation services. Founded in April 1996 by Brewster Kahle and Bruce Gilliat, Alexa Internet has harnessed the collective wisdom and experiences of the entire Internet community and transformed them into the Alexa service. The company donates a copy of the Web on an ongoing basis to the non-profit Internet Archive, which is endowed to preserve our digital heritage for scholarly access. Alexa Internet is located at The Presidio of San Francisco, PO Box 29141, San Francisco, California 94129-0141. For business information call 415/561-6900, fax 415/561-6795. The information email address is info@alexa.com. Contacts: Jennifer Weissman, Antenna Group, 415/977-1911, jennifer@antennagr.com, or Cynthia Lohr, Alexa Internet, 415/561-6786, cynthia@alexa.com.

Note to Editors: Alexa is a trademark of Alexa Internet. All other trademarks are the property of their respective owners.

CONTACT: Alexa Internet Cynthia Lohr, 415/561-6786

cynthia@alexa.com or Antenna Group Jennifer Weissman, 415/977-1911

jennifer@antennagr.com

KEYWORD: CALIFORNIA

INDUSTRY KEYWORD: COMPUTERS/ELECTRONICS COMED

INTERACTIVE/MULTIMEDIA/INTERNET

* Search the archives of the Los Angeles Times for similar stories. You will not be charged to look for stories, only to retrieve one.

Copyright Los Angeles Times

[PREV STORY](#)

[NEXT STORY](#)

Business	GO	Site Index	GO
----------	----	------------	----



PREVIOUS MATCH



SEARCH RESULTS



SEARCH PANEL



NEXT MATCH

SEARCHING FOR INFORMATION

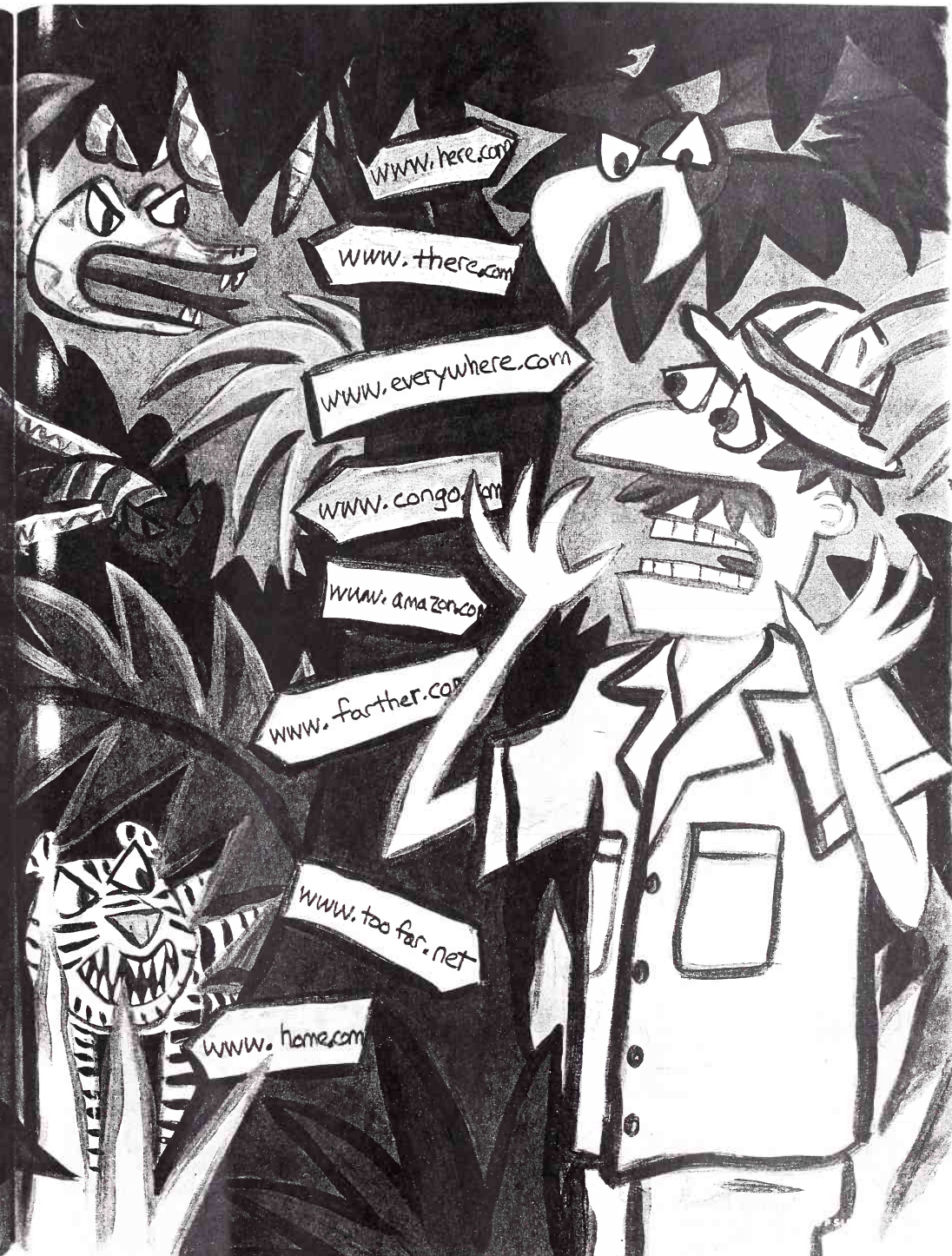
It's A Jungle Out There

BY JEFF UBOIS

Monday morning—your first day back at work after a vacation in the tropics. The trip was fabulous, but what is that nasty rash under your arm? Trying to avoid a trip to the doctor you hop on the Web. Giving AltaVista a whirl, you type in “jungle rot.” The service goes on safari in the digital jungle, and then ... argh, information overload. Alta Vista has found 39,741 references to jungle rot. While you scratch uncontrollably, you lament, “Can’t anyone come up with a better way to sort and retrieve information?”

ILLUSTRATION BY KIM WILSON BRANDT





www.here.com

www.there.com

www.everywhere.com

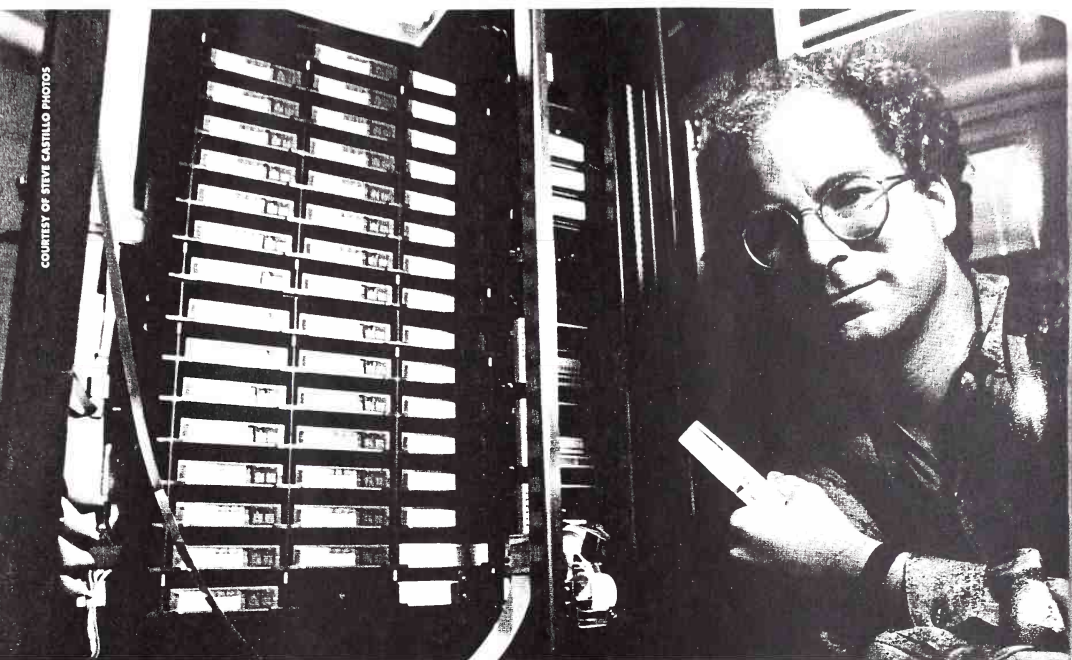
www.congo.com

www.amazon.com

www.farther.com

www.too far.net

www.home.com



Brewster Kahle, president of Alexa Internet, San Francisco

Fifty years ago, presidential science adviser Vannevar Bush accurately predicted the rise of personal computers, information storage and search tools, and electronic commerce in the article "As We May Think," published in the July 1945 issue of *The Atlantic Monthly*. It has become one of the most widely cited scientific articles of the century.

Bush, who oversaw all U.S. wartime science efforts (including the Manhattan Project), worried that increasing specialization would soon stall scientific progress. Noting that "Mendel's concept of the laws of genetics was lost to the world for a generation because his publication did not reach the few who were capable of grasping and extending it," Bush proposed a solution called the Memex. He envisioned "a device in which an individual stores all his books, records and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility."

It was a great vision, but the promise of the Memex remains unfulfilled. While Apple Computer Inc. triumphantly announced it had completed the Memex

when it shipped HyperCard more than a decade ago, the modern approximation of the device is the sum of Internet search tools and navigation services.

**Punch in "travel"
on Digital Equipment
Corp.'s AltaVista
search engine, and
you'll get more than
7 million possible
answers. That's noise,
not information.**

Companies offering these services have created billions of dollars in new market capitalization during the past two years

and are the core business of close to a fifth of the public Internet companies. But these companies and market observers agree that information searching is still harder and less comprehensive than it should be. So Internet startups and more established companies are investing millions of dollars in search research. These companies are pursuing a variety of strategies, including systems that provide visual representations of data, collaborative filters that gather recommendations from users, enterprise software enhanced with search capabilities and advanced Web-search engines.

A Difficult Problem

Search services have unified what was once a sprawl of uncharted fiefdoms on the Internet. But despite advances in processing and storage, the underlying approach has remained essentially unchanged for years: Users enter a few descriptive words and hope those words are used in relevant documents. It's fine for searching through a collection of a few thousand items, but punch in "travel" on Digital Equipment Corp.'s AltaVista search engine, and you'll

Jungle

get more than 7 million possible answers. That's noise, not information.

The explosion of data available on the Web is aggravating the problem. By late 1997, there were more than 640,000 sites and 100 million pages on the Web, and the number continues to double every six months. Without fundamentally new approaches, the number of documents returned by a generic search engine will grow at the same rate.

It's already causing changes in the way browsers operate. "We're looking at how to tightly couple the two main methods for information retrieval—search and browsing," says Susan Dumais, a guru in the field who was recently hired away from Bellcore to become a senior researcher at Microsoft Corp. But the real issue is deeper. The rapidly widening gap between the amount of data in the world and the

amount of attention available to process it means a growing percentage will never be looked at by a human. So the importance of search and filtering technologies will increase, even as new media types expand the challenge beyond the realm of straight text.

Searching for a Solution

"People have spent many millions of dollars trying to get traditional search technology to work," says Brewster Kahle, president of San Francisco-based Alexa Internet. "But as the amount of information and the number of people using it goes up, these traditional approaches are failing, and we need a new genre of answer."

Several dozen companies believe they've found it. Though they take an amazing variety of forms, business models and technical approaches, they all promise to determine the relevance of information for their users. Generally, they fall into three camps: free, ad-supported gateways to the Web; technical innovators; and enterprise-software companies.

A few of them will win big, as Yahoo Inc. already has. But most are fated to be acquired by big companies if things go well, or to close their doors. Search, filtering, personalization and agents—or whatever technologies eventually dominate—are too important to remain in the hands of small companies. Big companies need to own the best version of search, and they can afford to buy it.

It's not hard to imagine the commercial potential of the ultimate search tool or a magic relevance detector. Terabytes of unstructured information could be ordered and made accessible, duplication of effort would be reduced or eliminated, and companies could begin to get a return on unused information assets. And if the magic relevance detector were a channel of some kind, tuned perfectly to the user, why would anyone ever want to switch it off?

Portals and Gateways

Start with that last question first. Providing a portal or gateway to other sites rather than generating original content made Yahoo and AltaVista immediate successes and is the core business of Infoseek Corp. of Sunnyvale, Calif.; Excite Inc. of Redwood City, Calif.; Lycos Inc. of Framingham, Mass.; Inktomi Corp. of San Mateo, Calif.; and others. But anyone can be a gateway, and several of the early Web indices have either been shut down or acquired.

To create a sustainable model, Web directories and search engines are evolving into media brands that create destinations rather than serve as gateways. Their new competition is not just America Online Inc., but television and newspapers.

"People have to realize the Web has grown so much since the search days, and there is a need to expand beyond that one function," says Jerry Yang, co-founder of Yahoo. "What we do is a media game, not a software game, and search is only a complement to the overall service, not the dominant element."

George Bell, president and CEO of Excite, echoes these sentiments. "Technology enables everything we do, so it is enormously important, but this business is about building and growing brands," Bell says. "I'm less curious about the search market and more curious about how to use the search technology to bring a more useful and convenient experience." Innovators in the Web-wide search business are trying to improve results by providing better categorization of search results, automatic suggestions of new terms, improved "refine" features that make it easy to try again and the use of personalized pages that are essentially saved searches.

Elite Content

Rather than make searches more exhaustive, some believe the quickest way to improve the quality of search results is to index only top-quality content. "It's strange that a lot of the differentiation in the search engines is 'I index the most documents,'" says Steve Krause, president of Affinicast Corp. of San Francisco, which makes server software that matches reader interest with ads.

One company working with premium content is Northern Light Technology LLC of Cambridge, Mass., which has combined

continued on page 146

**It's not hard to imagine
the commercial potential of
the ultimate search tool or
a magic relevance detector.**



Jerry Yang, co-founder of Yahoo Inc., Santa Clara, Calif.

Jungle

Continued from page 110

a Web-wide search engine with automatic categorization. Traffic on the service (www.nlsrch.com) is growing at 20 percent per week, with 80,000 to 90,000 searches per day as of late December 1997. "We are a search engine, but we have this unique classification scheme we apply to results, and we offer access to premium material," says marketing manager Leslie Ray. "A lot of high-quality information from the publishing world on services such as [Reed Elsevier Inc.'s] Lexis-Nexis or [Dialog Corp.'s] Dialog has been inaccessible to the general public."

Excite is doing something similar, indexing 300 to 400 content sites three to four times a day. Bell says it means "80 percent of the value you get from Lexis-Nexis you get from us for free."

While limiting the collection of indexed documents may provide an interim solution, others are pursuing fundamentally new technologies.

Visualizing New Categories

Three classic science fiction stories about the Net—*True Names* and *Other Dangers* (1981) by Vernor Vinge, *Neuromancer* (1984) by William Gibson (which coined the word "cyberspace") and *Snow Crash* by Neal Stephenson (1993)—all describe user interfaces based on 3D representations of large data networks.

No one has yet built anything that matches the richness of the environments described in those literary works, but companies such as Lucent Technologies Inc.'s Visual Insights; ThemeMedia Inc. of Redmond, Wash.; Perspecta Inc. of San Francisco; Semio Corp. of San Mateo, Calif.; and Xerox Corp.'s Inxight are trying to tap into native human capacities to deal with 3D space. Instead of answering queries with lists of documents, these companies are building multidimensional, immersive environments that provide a more intuitive view of large collections of data grouped or clustered by meaning.

What's immediately striking about these offerings is their visual presentation. A 3D data representation can look like anything from a topographic map to a bike wheel with 100 spokes. But behind the pretty pictures,

they all rely on proprietary methods of categorizing documents and search results. Just as categorization makes the difference between a well-ordered library and a random heap of books, the value of the pictures these companies provide depends on their ability to group search results into meaningful categories.

A newcomer in this area is Visual Insights, an independent unit of Lucent, which unveiled a new set of products in December for visualizing semistructured data as well as information generated by automated systems. "There has been an explosion in the ability to collect data, but not to analyze it or take action," says James Weichel, president of the company. "A more challenging problem is visualizing data without a physical or geographic metaphor."

Weichel offers a striking example. Imagine trying to numerically describe a children's game in which players alternately choose numbers from one to nine in an attempt to get three of them that add up to 15. Each number may only be chosen once during the game, so player A might start by choosing seven and player B by choosing eight; player A then chooses six, and then player B would have to choose two; player A would respond with 5 to prevent B from winning.

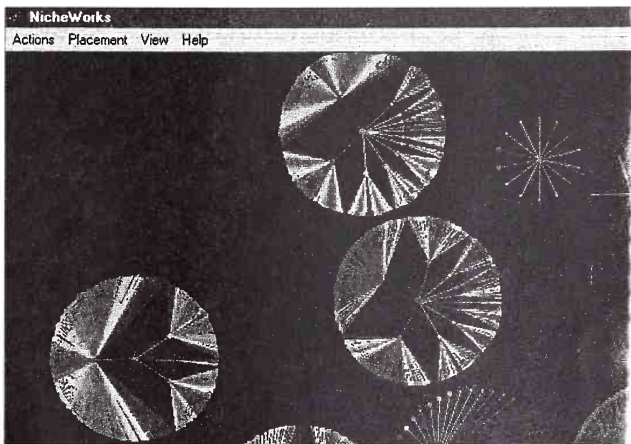
The variations are numerous and hard to solve—until you realize this scenario corresponds exactly to tic-tac-toe.

Weichel says similar leaps from data to

A 3D data representation can look like anything from a topographic map to a bike wheel with 100 spokes.

pictures are equally useful in areas such as network management, Year 2000 problem analysis, data mining, call-center operation and fraud detection. Much of the software and intellectual property now owned by Visual Insights (the company has nine patents) has been under development for several years, and the company, which employs about 40 people (more than any of the other visualization companies), is likely to eventually leave the shelter of Lucent through an IPO.

Like others in the search industry (such as Verity Inc. of Sunnyvale, Calif., and Aptex Software Inc. of San Diego), ThemeMedia has its roots in the U.S. intelligence community. Founded by researchers from the Pacific Northwest National Laboratory, ThemeMedia is led by supercomputer guru



How big is your company's year 2000 problem? In the Visual Insights screen grab above, each wheel represents a software system, such as payroll. Spokes represent individual programs—red spokes indicate programs with Year 2000 bugs. By clicking on the end of a spoke, a programmer can view and repair the actual code.

Jungle

Gary Smaby, who serves as CEO, and former Cray Research CEO John Rollwagen, who is chairman of the board.

"The best cognitive processor is between your ears, and a visual metaphor means that trends are patterns that can be seen," says Steve Ardire, ThemeMedia's senior director for business development. Though it isn't shipping any products yet, ThemeMedia plans to make money selling server software and providing free client software for viewing. "We have a core offering that could be a front end for an Intoseek or a collaborative filtering company such as Net Perceptions or Autonomy," Ardire says.

The hope is to go beyond the limits of Boolean searching, which is based on particular terms and logical operators such as AND, OR and NOT. "The problem with Boolean is that users have to know specific facts about documents before they conduct a search," Ardire says. "[Our technology] lets you drill down deeper to reveal new data, and none of the Boolean offerings allow you to do that."

Another company using visualization is Perspecta. "If you know exactly how to define a query, a search engine is useful," says President and CEO Steve Holtzman. "We

are aiming at the other 90 percent who can't define exactly what they want, and we allow them to dynamically reorganize the data depending on how they want to see it" (for example, reordering a news database according to date, topic or source).

Founded in January 1996 by MIT Media Lab alumni, Perspecta has received funding from Encyclopaedia Britannica and Informix Corp. and counts these companies, as well as The Sabre Group, as customers.

Visualization-software companies share the ability to present an immediate and intuitive connection between different clusters of data and to relate those to a user's position in a 3D environment. But it's not the only new approach.



George Bell, president and CEO of Excite Inc., Redwood City, Calif.

Collaborative Filters

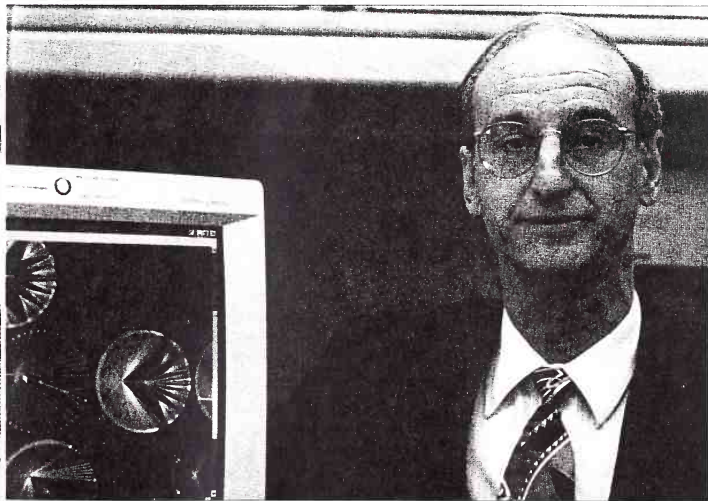
How do you know what movie to see, what restaurant to eat at, or what book to read? If you're like most people, you rely on the recommendations of peers. The Net offers a unique way for individuals to pool their collective wisdom through collaborative filters, or recommendation systems. A collaborative filter can allow a Web site to offer restaurant ratings, for example. When people visit the site, they are asked to rate a few restaurants, and the system then finds others with similar preferences and makes suggestions about places to eat.

Collaborative filters vary by the type of subject matter they deal with, whether they match people to people or people to things, and whether they collect data actively and explicitly, passively and implicitly, or in combination. But they all provide answers to questions and predict what a given person will like by condensing opinions from thousands of users. Unlike search engines, collaborative filters learn about the preferences of individual users, and they can offer recommendations about things that can't be indexed, such as video, for example.

There are a several companies and products in this space, including Net Perceptions Inc.'s GroupLens Recommendation Engine, which helps companies match individual customers' tastes and is backed by The Paul Allen Group and Hammer Winblad Venture Partners; Autonomy Inc.'s Agentware i3, which helps Web publishers deliver targeted content; LikeMinds Inc. of San Francisco, which focuses on one-to-one marketing and has a movie recommendation demo; Gustos Software LLC of Laguna Hills, Calif., which lets users rate Web sites; and Pittsburgh-based WiseWire Corp.'s WiseWire information service.

But the most prominent is Firefly Network Inc., a Cambridge, Mass.-based company founded in 1995 by another group of MIT Media Lab researchers. It was first out of the gate with a music recommendation system (later sold to a startup called Launch). "We are moving toward a model of computing and information retrieval that is based around the user instead of an operating system or datacenter model," says Saul Klein, Firefly's senior VP of corporate brand and strategy.

Firefly has fallen short of some of its early, perhaps overinflated, hopes—the company was supposed to be public by now but isn't.



James Weichel, president of Visual Insights, a unit of Lucent Technologies.

Jungle

In addition, the math behind collaborative filtering is easy to duplicate, which has split the market among many vendors, and the killer app (or the killer data set) for collaborative filtering has been elusive.

"I don't want to knock collaborative filtering—we use it," says Yang of Yahoo, which is using Firefly in its My Yahoo service. "Firefly has done a good job with movies and music, but our focus now is to look at such technology to improve our content targeting and advertising targeting." Striking a balance between active/explicit and passive/implicit search and navigation hasn't been done. It's an area ripe for new solutions, and one of the more intriguing is Alexa Internet.

The Alexa service is delivered via a client that looks like an additional toolbar for the Netscape or Internet Explorer browsers. Alexa displays information about whatever site a user is accessing, lets users vote on site quality and offers suggestions about where to go next.

Alexa's information is generated by end users, from third-party information providers (including Encyclopaedia Britannica) and from an analysis of the company's copy of the entire Web. By gathering copies of all publicly accessible sites (a collection that is now 8 terabytes and growing), Alexa can analyze the link structure of the Web to determine which pages are relevant to each other, and allow users to request archived Web pages when a "404—File Not Found" error occurs. Data gathered by Alexa is donated to the nonprofit Internet Archive (www.archive.org) for long-term safekeeping.

"If you take a path through the woods, you benefit from the exploration people have done before you in finding the best way up the mountain or down to the lake," says Alexa's Kahle. "We do the same thing for the Web. When your browser hits a URL, the toolbar requests information about it

from Alexa's servers, and we record that an Alexa user has spent time at that site as a kind of vote."

What Alexa has in common with the collaborative filtering and recommendation systems is an approach based on metadata, a loosely defined term that is rapidly becoming a buzzword used to describe any data about other data. A call number in a library, a keyword describing an article

and author tags in HTML are all examples of metadata.

Metadata is becoming increasingly important because it enables searches through content that can't be indexed easily, as well as through enormous collections of data. Metadata has become a major focus for the World Wide Web Consortium, which is developing a new standard called XML. This is a specification for defining new markup languages to meet the future needs of Web documents and Web searching.

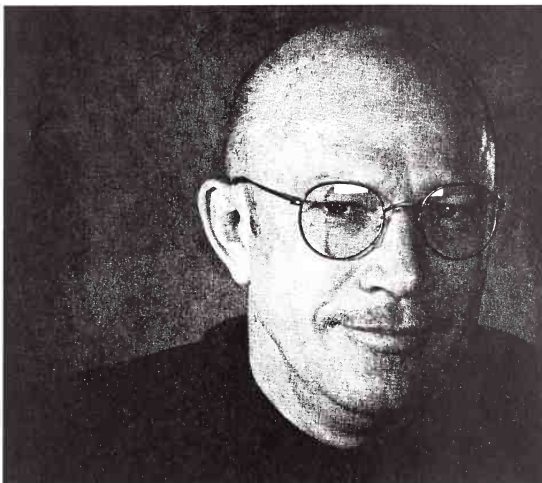
Enterprise Software

Companies such as Verity; Fulcrum Technologies Inc. of Ottawa, Ontario; CompassWare Development Inc. of New York; Quadralay Corp. of Austin, Texas; Personal Library Software (PLS) Inc. of Rockville, Md.; Open Text Corp. of Waterloo, Ontario; Open Market Inc.'s Folio unit; and Excalibur Technologies Corp. of Vienna, Va., are pursuing opportunities created by corporate intranets. They're competing with offerings from the public search engine companies, including AltaVista, Infoseek, Lycos and Excite, all of which

are licensing the software they use to support their public sites.

Like the "elite content" companies, these firms know that the value of the data being searched determines the value of search applications. And because data inside companies is extremely valuable, the tools that access this data are valuable, too.

One of the most established companies in this space is Verity, which emphasizes the work it has done to build interfaces to other vendors' database products and is hoping to distance itself from low-end search products. With dozens of free or low-cost search tools available, Verity wants to stand out from the crowd by building connections into Lotus Notes and other large databases, not by adding a more sophisticated user interface. "Search and retrieval is fragmenting into two or more primary camps: commodity personal search and tools for corporate use," says Ronald Weissman, Verity's VP of marketing.



Brett Newbold, president, Open Text Corp., Waterloo, Ontario.

"If you take a path through the woods, you benefit from the exploration people have done before you. ... We do the same thing for the Web," says Alexa's Brewster Kahle.

Jungle

Commoditization is an issue, but the most serious threat to the small and midsize search companies is competition from IBM Corp., Oracle Corp. and Microsoft, all of which are adding functions to existing products and developing new ones.

The stock prices tell the story. By the end of 1997, Fulcrum had dropped from a 1996 high of \$20.50 to less than \$2, Verity had plunged from more than \$50 in spring 1996 to about \$5, and Excilibur had recovered to about half the value it had in early 1997.

"Verity, PLS, Fulcrum and similar companies are having a hard time because search by itself doesn't solve anyone's real problem. There is much more to do," says Open Text president Brett Newbold. "How do you compete with Oracle when Oracle says it is free, or Microsoft makes it a free part of Exchange?" Smaller companies are running for cover by creating focused applications. "Every [company that] was anywhere near this space is now calling itself a 'knowledge management' company," Newbold says. Knowledge management is evidently on its way to becoming the total quality management or re-engineering of the late 1990s—even Open Text is describing itself as a "collaborative knowledge management company."

Customer service is another good candidate application in this area. Paul McGraw, executive VP and co-owner of APS Technologies, a Kansas City, Mo., supplier of SCSI and other peripheral products, says his company would like to find a way to categorize and access the knowledge contained in the problem-resolution notes taken by technical support representatives. APS has used an extensive coding scheme to key into typical problems, but tapping the megabytes of text notes is a daunting task.

"A simple search may not find parallel port cables with SCSI, or a machine that freezes may not be diagnosed correctly as having a SCSI problem," McGraw says. APS has worked with software from Folio to provide some context-sensitive help, but its greatest successes in diagnosing problems have come from manually analyzing the data, or from running SQL queries against an AS/400 database.

Clearly, there is room for improvement

and a market for enterprise tools. But advances in this area will depend on the evolution of new search technology.

Future Search

Yahoo has shown that ads can support navigation services and that branding can be as important as technology. If there are no great leaps forward in search technology, refining the art of human assistance will be the next-best thing, and the art of building search services will depend on knowing where to add human help, as Yahoo has done.

"We don't believe there is this magic technology to make things much better or worse," Yang says. "Anything we adopt has to address 80 percent of the users. If it's useful for only a few, it's not our business and we can't support it."

For any technology player that hopes to win, Yang adds, an offering "not only has to scale, but also must be compelling enough that lots of people want to use it, and most have failed at either or both." Still, Digital proved the importance of technical brute

Commoditization is an issue, but the most serious threat to the small and midsize search companies is competition from IBM Corp., Oracle Corp. and Microsoft.

force when its AltaVista service increased the value of the Web by an order of magnitude overnight. New ways to cluster and view search results and to iterate searches could do the same. Standards to allow

continued on page 155

Digging for Data: Where and How to Look

Some of the frustration with search engines can be traced to the fact that most users don't have a clue how to use the technology effectively. For starters, entering just one word into a search engine is like using a backhoe to pick out a splinter. Search engine operators say that among those rare users who enter more than a single term, few think carefully about synonyms or are adept at making queries using Boolean logic—for example, "A AND (B OR C)." Among those who do understand Boolean logic, many don't take the time to read the directions for each search engine's command syntax.

But because search technologies are still stupid, users have to be smarter about making them work and about zeroing in on the good stuff by using several engines. One option is to use metasearch services—such as www.mamma.com, www.melosearch.com and www.search.com—which combine results from more than one database. Search experts also suggest that those digging for data focus their search re-

quests, find unique identifiers associated with the topic (proper names, geography, dates and jargon), play with synonyms and keep trying.

The best searches work by iteration, and sometimes by coming at the question from a different angle or by using a different set of data (Usenet archives rather than the Web, for example). Maybe the right answer isn't indexed on the Web but can be found within a vendor's Web site. The trick is to know how and when to broaden a search and when to narrow it.

You may also want to consider using a fee-based service. Offerings include Northern Light Technology's service (www.nlsearch.com) and Infonautics' Electric Library (www.elibrary.com). Or if you want to pay through the nose, try Dialog Corp.'s Dialog service (www.dialog.com) or Reed Elsevier Inc.'s Lexis-Nexis (www.lexis-nexis.com). These last two services charge for searches based on the libraries accessed, the connect time required and the number of documents printed. —J.U.

Search Sectors

Notable players in four categories of software

The only way to categorize the companies developing search technology is to be arbitrary about it. The boundaries between different segments are blurry (Excite, Infoseek and other Web indices sell their software for use on intranets), and every successful new company will end up competing against—or possibly being acquired by—the likes of Microsoft, IBM, Oracle and Netscape.

—Michael Mattis and Jeff Ubois

Traditional, Ad-supported Search Engines and Directories

"In terms of aggregate traffic, ad revenue and brand recognition, Yahoo is clearly the leader in the search and directory space. For a long time it has been a two-horse race with Yahoo still well ahead of Excite. This space is no longer about search, as the technology of search has essentially become a commodity (witness the fact that Yahoo even licenses search technology from AltaVista). Instead of trying to implicitly compete with one another, search and directory companies are starting to aggressively pursue the same audience as AOL—they are competing with online services (as news and general information sites). AOL has even sold its equity stake in Excite, as they now see the company as a competitor."

—Patrick Keane, Jupiter Communications LLP

Yahoo Inc.

Location: Santa Clara, Calif.
URL: www.yahoo.com
Ownership: Public (Nasdaq:YHOO)
Main Product: Yahoo (Yet Another Hierarchical Official Oracle) software locates, identifies and sorts material stored on the Internet.

Excite Inc.

Location: Redwood City, Calif.
URL: www.excite.com
Ownership: Public (Nasdaq:XCIT). (Inuit Inc. recently invested \$40 million in Excite.)
Main Product: Excite ICE (Intelligent Concept Extraction) searches the Web and Usenet for documents containing keywords and related concepts.

Infoseek Corp.

Location: Sunnyvale, Calif.
URL: www.infoseek.com
Ownership: Public (Nasdaq:SEEO)
Main Product: Infoseek Search searches the Web, specific sites, URLs, titles or links without Boolean operators.

Lycos Inc.

Location: Framingham, Mass.
URL: www.lycos.com
Ownership: Public (Nasdaq:LCOS)
Main Product: Lycos Search locates, retrieves and manages information stored on the Net.

Inktomi Corp.

Location: San Mateo, Calif.
URL: www.inktomi.com
Ownership: Private
Main Product: Inktomi Search Technology is a scalable engine for searching and retrieving information stored on the Internet. It powers such services as Wired's HotBot.

AltaVista Search Network Inc.

Location: Palo Alto, Calif.
URL: <http://altavista.digital.com>
Ownership: Digital Equipment Corp. (NYSE:DEC)
Main Product: AltaVista Public Search aggregates and retrieves data from the Internet. AltaVista licensing includes deals with TPI (Telefonica Publicidad e Informacion, Spain) and Yahoo Inc.

Recommendation Systems/Collaborative Filtering

"The players in this emerging market consist of technology enablers, which are attempting to create one-to-one marketing solutions, and full-service solutions providers, which are not only creating tools, but also are creating the platforms to allow content and commerce companies to take personalization and its opportunities to the next level.

"In this pack, Firefly has the best vision of an end-to-end solution. Since its inception, the company has shown how its product can make communities come alive in a number of settings. In particular, Firefly has shone in a number of music and media-rich applications. [Among] others that have the guile and experience to stand out [is] Alexa, the brainchild of the folks behind WAIS and NetPerceptions."

—Allen Weiner, Dataquest Inc.

Firefly Network Inc.

Location: Cambridge, Mass.
URL: www.firefly.net
Ownership: Private
Main Product: Firefly Catalog Navigator personalization software captures and adds preference and general interest-level information about client businesses and their products according to customers' personal profiles.

Net Perceptions Inc.

Location: Minneapolis
URL: www.netperceptions.com
Ownership: Private
Main Product: GroupLens Recommendation Engine tailors a client Web site's content to each of the site's customers.

LikeMinds Inc.

Location: San Francisco
URL: www.likeminds.com
Ownership: Private
Main Product: WebSell is a collaborative filtering and predictive modeling technology that taps customer's behavior data based on their online interactions. It then identifies product selections that have the "best chance" of being purchased by targeted customers.

Alexa Internet

Location: San Francisco
URL: www.alexa.com
Ownership: Private
Main Product: Alexa Internet is delivered via a client that looks like an additional toolbar for the Netscape or Internet Explorer browsers. Alexa displays information about the site a user is accessing, lets users vote on site quality and offers suggestions about where to go next based on the sites this user, or others like him or her, have visited. Alexa's information is generated by end users, from third-party information providers, including the Encyclopaedia Britannica, and from an analysis of the company's copy of the entire Web.

Imana Inc.

Location: San Francisco
URL: www.imana.com
Ownership: Private
Main Product: Using agent technology, Imana's Talkman software enables product and service companies to offer personalized information delivery, targeted advertising and self-forming online communities.

WiseWire Corp.

Location: Pittsburgh, Pa.
URL: www.wisewire.com
Ownership: Private
Main Product: Using what the company calls smart filtering, WiseWire for Web Sites integrates into a business's Web site, linking it to updated, real-time information specific to a business's objectives in order to build community around shared interests.

Gustos Software LLC

Location: Laguna Hills, Calif.
URL: www.gustos.com
Ownership: Private
Main Product: Gustos Guide is a recommendation engine for commerce and ad servers. Using collaborative filtering, it lets users rate Web sites.

Jungle

Continued from page 152

Enterprise/Intranet Software

"While Verity raced to an early lead as the premier search engine for enterprise Web sites, it is now facing stiff competition from multiple sources. Moreover, with the advent of Microsoft's Index Server as a core component of a free NT Web server, basic full-text retrieval is rapidly becoming a commodity service in cyberspace. Needed are network-centric solutions that offer superior retrieval and recall, with a minimum of administrative overhead."

—Geoffrey Bock,
Patricia Seybold Group

Verity Inc.

Location: Sunnyvale, Calif.
URL: www.verity.com
Ownership: Public
(Nasdaq: VRTY)
Main Product: Search'97 is a scalable, customizable knowledge retrieval solution for organizing and navigating enterprise information resources.

Microsoft Corp.

Location: Redmond, Wash.
URL: www.microsoft.com
Ownership: Public
(Nasdaq: MSFT)
Main Product: Microsoft Index Server is a full-text retrieval engine that comes included with Microsoft's NT Web server software.

Excalibur

Technologies Corp.

Location: Vienna, Va.
URL: www.excalib.com
Ownership: Public
(Nasdaq: EXCA)
Main Product: Excalibur RetrievalWare is a set of tools for building text-based knowledge retrieval solutions that can search through information located on corporate networks, including intranets, the Web and other knowledge repositories.

Fulcrum

Technologies Inc.

Location: Ottawa, Ontario, Canada
URL: www.fulcrum.com
Ownership: Public
(Nasdaq: FULCF)
Main Product: Fulcrum Knowledge Network is an information retrieval solution for business. It provides agents, document summarization and intuitive searches of the Web, file servers, intranets and databases.

Open Text Corp.

Location: Waterloo, Ontario
URL: www.opentext.com
Ownership: Public
(Nasdaq: OTEKF)
Main Product: LiveLink Intranet is a Web-based, scalable, collaborative knowledge management application for intranets. A component of the technology, LiveLink Spider, crawls designated intranet and Internet Web sites, retrieving data that is then indexed and catalogued.

Perspecta Inc.

Location: San Francisco
URL: www.perspecta.com
Ownership: Private
Main Product: The SmartContent System is a scalable development platform for navigation, online analysis and profiling for applications using 3D visual interfaces. It facilitates interaction between a company and its customers, distributors, suppliers and partners.

Inxight

Software Inc.

Location: Palo Alto, Calif.
URL: www.inxight.com
Ownership: A Xerox Corp. enterprise company (NYSE:XXR)
Main Product: VizControls displays large quantities of information as graphic or visual representations. Users view thousands of data points, documents or information objects at once in various graphical ways.

Personal Library Software (PLS) Inc.

Location: Rockville, Md.
URL: www.pls.com
Ownership: Joint venture Netscape and GE Information Services
Main Product: The PLS family combines enterprise administration with full-text information retrieval on corporate intranets and the Web.

Sovereign Hill Software Inc.

Location: Hadley, Mass.
URL: www.sovereign-hill.com
Ownership: Private
Main Product: Sovereign Hill's InQuery distributed intelligent search accepts queries in plain English and then finds, retrieves and sorts data from appropriate databases.

searching through multiple databases will also be important. A truly robust software standard for database access could shift the battleground from the Web back to the desktop-software market.

At the same time, search is becoming a commodity; its future value depends increasingly on the associated data and applications. Poor search technology married to excellent content or applications beats great search technology applied to trivia.

For investors, that means search companies without solid applications are, or should be, passé. The likely winners will be the big companies that traditionally take over commodity businesses and small companies that marry search to more focused applications.

Many search tools could follow the path of desktop publishing to become a feature in every application rather than something that stands alone. "We are going to see search functionality built into many applications," predicts Microsoft's Dumais.

Though the early winners will be those that make the fewest demands on end users, search is important enough that users will change their behavior to get better results. And as Vannevar Bush indicated, it will change how people think, offering the freedom to forget, to remember and to know at will.

"[Man's] excursion may be more enjoyable if he can reacquire the privilege of forgetting the manifold things he does not need to have immediately at hand," Bush wrote, "with some assurance he can find them again if they prove important." ■

Jeff Ubois (jubo@netcom.com) is an independent consultant who has been digging through large databases and writing about the Internet for more than a decade.

3D Visualization

"Data visualization is the latest emerging trend for analysis of large volumes of detailed data. Current grid and chart presentation methods are too limited for viewing complex result sets with many different variables. As an adjunct for reporting and data mining, visualization tools give the end user the ability to comprehend tens of thousands of items in a single view. We see visualization tools and components as a growing market over the next two years. [However,] it is too early to see a clear front-runner in this emerging sector."

—Dan MacLovich, Meta Group Inc.

Visual Insights

Location: Naperville, Ill.
URL: www.lucent.com/visualinsights
Ownership: A unit of Lucent Technologies Inc. (NYSE: LU)
Main Product: Visual Insights is interactive data visualization software developed at Bell Labs. It displays 3D macroscopic views of data that can be manipulated by users.

ThemeMedia Inc.

Location: Redmond, Wash.
URL: www.themedia.com
Ownership: Private
Main Product: SPIRIX is a prototype intended to replace text-based retrieval and analysis with an interactive, 3D visualization-based system that uses landscape-like interfaces to categorize data.



Microsoft **Daily News**
7/18/97

[Click here to get the full story](#)



**American Express and Microsoft®
team up to trounce travel costs**

INTERNET
MAGAZINE

[Demos & Downloads](#) [ZD Labs](#) [Editorial & Opinions](#)
[Departments A-Z](#) [Website Makeovers](#)



NEWSwatch

FOR FRIDAY, JULY 18, 1997

Alexa Debuts Search Engine That Learns from Its Users

[NEWSwatch Archives](#)

[Alexa Internet](#) yesterday introduced a new search engine that is designed to learn from its users, providing Website recommendations based on analysis of anonymous Internet usage patterns.

Working in conjunction with standard Web browsers, the Alexa navigation service resides on the browser as a toolbar displayed at the bottom of a page. The toolbar lets users know where they are on the Web, providing detailed Website information including who the site is registered to, how many pages the site hosts, how frequently the site is updated, and how popular the site is with others.

The toolbar also suggests other sites to surf to next, and alerts users as to whether colleagues and friends are online, enabling real-time online communication.

The technology behind Alexa is based on the overall traffic patterns of the Web, the ecology of links between sites, the content of Web pages, and anonymous usage paths.

Available for Windows 95 and NT, an Alexa beta version is currently available. General availability of the free service is projected for late summer.

NEWSwatch stories for Friday, July 18, 1997:

- [Centra Unveils Online Learning, Collaboration Tool](#)
- [Company Offers Network Design Tool for LANs, WANs](#)
- [Alexa Debuts Search Engine That Learns from Its Users](#)
- [SureSite Announces Free Web Page Builder](#)
- [Crystal Graphics Uncorks 3D Graphics Development Tool](#)
- [NCSA Forms ISP Security Consortium](#)



TOP



E-MAIL

— A ZDNet Site —

[ZDNET HOME](#)

[SITE GUIDE](#)

[FIND IT](#)

[COMMUNITY](#)

[AD INFO](#)

[CONTACT US](#)

Good Morning Silicon Valley

Only in Mercury Center

EDITION

TIME (PST)

First Light

Morning

Tech Stocks

Tech Ticker

Archive

Breaking News

About this site

© 1997 Mercury Center
All rights reserved.

Mercury Center Home Page Feedback
This Mercury Center information you receive
is protected by the copyright laws of the United States. The copyright
prohibits any copying, retransmitting,
posting, or otherwise using any
copyright-protected material.

MERCURY CENTER SAN JOSE MERCURY NEWS

Morning

Last updated: Friday, July 18, 1997, 8:30 a.m.

Frames: [Disable](#) / [Enable](#)

You'll see other explanations, but the real reason that [web](#)
[convulsed](#) yesterday? We've got a full moon.

Firing Transfer Protocol: FTP Software of Andover, Mass.
has [laid off 300](#) of its 300 employees because of continuing losses.

Deep pockets: Microsoft's [profits rose 89 percent](#) in the past
quarter, and company spokesmen are already trying to lower
expectations for the future.

Bio-tech warriors: A [new class of antibiotics](#) has been created
by researchers at Stanford University and Kosan Biosciences.
They genetically altered the popular antibiotic erythromycin to
arrive at the discovery.

Ban phone tag, too: The U.S. Senate has decided [no more](#)
[computer games](#) for government employees.

Feel like a trip? Microsoft is thinking about [overseas](#)
[expansion](#), particularly in Dublin, Ireland, or India. The high-tech
labor pool is growing shallow around Redmond (or in the U.S.,
for that matter), which is making Microsoft think global.

No-meddle zone: A Philadelphia city council member asked a
Congressional subcommittee to [let the locals tax the Net](#) if they so
desire. Also, a motion picture mogul wants more copyright
protection for creative works on the Net.

Pushing is hard: Netscape [pushes hard](#) for its Netcaster
software to beat out Microsoft's IE. There's another beta version
available now and the actual release is set for next month. (BTW,
there's a [bug fix](#) for Communicator that's also available now, too.)

Filmstrips of the '90s: For an [alternative view](#) of computers in
schools, the Atlantic Monthly casts a skeptical eye on the rush to
wire our schools -- especially at the expense of art, music and
physical education.

Cue the scary music: Sybase's Mitchell Kurtzman takes [dead](#)
[aim](#) at Oracle with the best quote of the morning: "You know in
those horror movies, where there's always a scene when people
think someone is dead, and then that person comes back to life and
scars the hell out of everybody? I think we're going to scare the
hell out of Oracle. They better watch out." (Another [interview with](#)
[Kurtzman](#) popped up recently in Steve Gillmor's column.)

Living in the future: The only sure thing about how the
technology world will look in the next five years is that it will be
[nothing like the present](#). Networking of homes is just a start.

Re-engineering surfing: A new development from a guy who
invented WAIS and was the lead engineer on the Thinking
Machines parallel-processing supercomputer project: Why not
[suggest like sites](#) for people via a browser companion application?

In Mercury
Center today:

- [Snafu sets off Net](#)
[chaos](#)
- [Wired's](#)
[co-founder leaves](#)
- [Lucent snaps up](#)
[Octel](#)

Stocks

The latest [stock and](#)
[market information](#) in
Mercury Center's stock
page



Get GMSV
Morning by
e-mail using
[Netscape's In-Box](#)
[Direct](#) (click on
"technology news")

By [Patricia Sullivan](#), Online Editor
Write to us at morning@sjmercury.com



Search
Site Map
What's New

Page One

Internet

Personal Computing

Corporate Computing

Business

Rumors & Comment

Headline Scan

AnchorDesk

ZDNN Radio

Personal News

Company Finder

News Talk

ZD Magazines

Get a
PC checkup!



?

ZDNet • **News** • Products • Web • Mac • Community
Magazines • Downloads • Games • At Home • Shopping • HealthyPC



Internet historian creates tool to inform Web site visitors

By Renee Deger
July 18, 1997 4:22 PM PDT
ZDNN

The Internet's historian now wants to be its guide. Brewster Kahle, the man who is attempting to archive every Web page for history, has just started a new venture that would inform Web page visitors all about the site they are visiting -- everything from who runs it to how many hits it records.

His Alexa Internet, a search engine now undergoing a limited public beta test, features a tool bar on the browser page that lists information about the Web site being viewed. Information includes who runs the site, how often it is refreshed and how much activity it sees. The tool bar also offers links to related sites, and gives details of who links to the site.

Such information would have come in handy to recent visitors of Nasa.com. Expecting news about Mars, users instead were greeted with a pornographic site. The NASA site is actually Nasa.gov, and Alexa Internet could have told users the other was owned and maintained by an entity other than the U.S. government.

The porn site is just a symptom of an emerging problem with cyberspace. Anyone can throw up a Web site and lure people in, and not all sites are so obviously revealed.

The source of Alexa Internet's data is the Internet Archive, a project begun by Kahle last year to save for posterity every one of the 50 million or so Web pages that exist. The cataloging information is now shared between Internet Archive and the 15-month-old Alexa Internet. Both are based in San Francisco.

The Net's historian himself has a long history with the Internet. To help the average user tap into the collective knowledge stored online, Kahle helped seed the online publishing industry in 1989. He founded WAIS Inc., an electronic publishing company. It's since been sold to America Online. He also was a co-founder of supercomputer maker Thinking Machines.

Now the critical mass of online information that he encouraged is now ready for a new navigation tool, one that

SEE

[Brewster's trillions](#)

[Letting Customers Dig Through Your Data](#)

[Some Mars seekers found porn instead](#)

offers qualitative analysis.

Kahle's site Thursday began registering beta users. He wants to enlist 10,000 users for the test, but believes the company ultimately can accommodate millions.

The search engine serves a dual role. It will tell users information about the site they are visiting, but it also records the visitor's travel habits. The program automatically communicates back to Alexa all of the sites its users visit and in what order, then aggregates the data. The results are incorporated into updates of the program, which are sent to users regularly and automatically, so users can view the traffic patterns even on new sites.

Kahle said the purpose of tracing user habits is so that people can learn about the Web from each other, albeit anonymously.

"We're not trying to profile who [users] are, we don't care. We want to aggregate the paths on the Internet," said Kahle.

The search engine will be free once it is finished later this year, said Kahle. The company, seeded with about \$5 million from a Swiss investment company, will rely on advertising for revenue.

The ads will rest in the toolbar and will change according to the kind of Web site the user is looking at. Kahle said he plans to forbid companies from buying spots that pop up on the toolbar when users visit a competitor's Web site.

Search 

Advanced Search



— A ZDNet Site —

[ZDNET HOME](#)

[SITE GUIDE](#)

[FIND IT](#)

[COMMUNITY](#)

[AD INFO](#)

[CONTACT US](#)

Copyright (c) 1997 ZDNet. All rights reserved. Reproduction in whole or in part in any form or medium without express written permission of ZDNet is prohibited. ZDNet and the ZDNet logo are trademarks of Ziff-Davis Publishing Company.



Microsoft® **Daily News**
7/18/97

Click here to get the full story



**American Express and Microsoft®
team up to trounce travel costs**

INTERNET
MAGAZINE

Demos & Downloads ZD Labs Editorial & Opinions
Departments A-Z Website Makeovers



NEWSwatch

F O R F R I D A Y , J U L Y 1 8 , 1 9 9 7

Alexa Debuts Search Engine That Learns from Its Users

[NEWSwatch Archives](#)

Alexa Internet yesterday introduced a new search engine that is designed to learn from its users, providing Website recommendations based on analysis of anonymous Internet usage patterns.

Working in conjunction with standard Web browsers, the Alexa navigation service resides on the browser as a toolbar displayed at the bottom of a page. The toolbar lets users know where they are on the Web, providing detailed Website information including who the site is registered to, how many pages the site hosts, how frequently the site is updated, and how popular the site is with others.

The toolbar also suggests other sites to surf to next, and alerts users as to whether colleagues and friends are online, enabling real-time online communication.

The technology behind Alexa is based on the overall traffic patterns of the Web, the ecology of links between sites, the content of Web pages, and anonymous usage paths.

Available for Windows 95 and NT, an Alexa beta version is currently available. General availability of the free service is projected for late summer.

NEWSwatch stories for Friday, July 18, 1997:

- [Centre Unveils Online Learning, Collaboration Tool](#)
- [Company Offers Network Design Tool for LANs, WANs](#)
- [Alexa Debuts Search Engine That Learns from Its Users](#)
- [SureSite Announces Free Web Page Builder](#)
- [Crystal Graphics Unveils 3D Graphics Development Tool](#)
- [NCSA Forms ISP Security Consortium](#)



TOP



E-MAIL

A ZDNet Site

[ZDNET HOME](#)

[SITE GUIDE](#)

[FIND IT](#)

[COMMUNITY](#)

[AD INFO](#)

[CONTACT US](#)

Alexa Browser Companion Makes Searches More Certain

by Lisa Moskowitz, PC World

July 18, 1997

If browsers only knew your train of thought when you scanned a site, life would be grand. No more clicking back and forth between search engines and sites, hoping that one of your 13,369 search results will offer up a sliver of useful information. Sigh no more, dear Netizens, Alexa has arrived.

Announced July 17, Alexa Internet is a free Web navigation service that learns from users' collective surfing patterns. Working independently of--but in conjunction with--your browser, Alexa provides recommendations on where to go next based on a site's content and the subsequent paths taken by previous visitors to the site.

For example, when you go to a site about camping, the Alexa toolbar lets the system know where you are so it can retrieve data about that site. In seconds, you'll receive facts about the site, like how many pages it has and to whom it's registered. You'll also be able to consult a list of related sites on subjects such as the best camping grounds in California or where to buy a tent. These suggestions pop up in a Where To Go section and are based on Alexa's user tracking. Subject-sensitive advertising, Alexa's revenue source, is also displayed in this section.

As you surf, your search path is sent anonymously to Alexa and incorporated into their database. The next user who comes to that site can then benefit from your search decisions. Privacy is not compromised because a rigid system discards the user's identity before any path details are forwarded, according to Alexa co-founder Bruce Gilliat.

If the browser can't locate a site for whatever reason, Alexa will pull up the most recent rendition of it from its archive. Alexa also has an instant chat feature for real-time communication with other Alexa users.

Alexa is the brainchild of Gilliat and Brewster Kahle, the meta-data king of Wide Area Information Server (WAIS) fame. Kahle invented the information retrieval system, which fetches documents by keywords and ranks them based on the number of times the keywords appears. Sound familiar? It should. Kahle sold WAIS to America Online in 1995. He and Gilliat, who worked at WAIS, then started on their new project: archiving the Web. "We thought we should preserve our digital heritage," Gilliat said. "The question then was, what can we do with this information to provide a service for Net users?"

The answer was Alexa. The start-up was founded in 1996 and accesses an archive containing more than 5TB (about 5000 gigabytes) of Web data. Alexa requires less than 10 percent of a user's bandwidth to operate, Gilliat said.

The beta version of Alexa is currently available for download on the Alexa home page for the first 10,000 registered users. The program works with Netscape and Microsoft browsers and requires Windows 95 or NT.

Internet Links

[IE 4: It's Here and It's Hot](#)

[Netscape Beta 3 Ready for Download](#)

[Who's On \(A\) First? Netscape Previews Netscaper](#)

[Netscape Beta 3 Ready for Download](#)

[MCI and Prodigy Dish Up New Net Access](#)

[String of Crashes Leads to Week From Hell for the Net](#)

[Netscape Beta 3 Ready for Download](#)

[More Daily News](#)

[Alexa Internet](#)

[America Online](#)

atOnce Software
World Shopper

Copyright © 1997 PC World Communications. All Rights Reserved. Use of this service is subject to the PC World Online Terms of Service Agreement.

Advertising
Sponsor - Please
Click Here

Introducing...

MORE SPORTS TITLES!
Buy 1 get 1 free!
Sierra

Click Here

The New York Times

Copyright © 1997 The New York Times

MONDAY, JULY 21, 1997

Printed

Technology

New Service Tracks Web Use

By JOHN MARKOFF

SAN FRANCISCO, July 19 — As a second generation of Internet search tools emerges, a company called Alexa Internet has started a service offering users more information about the popularity and location of the pages they view.

The service, which works with Internet browsers offered by the Netscape Communications Corporation and the Microsoft Corporation, focuses on the relatively new idea of assisting users in navigating the World Wide Web.

The Alexa service was founded by Brewster Kahle, who was one of the founders of the Thinking Machines Corporation, and by Wais Inc., an information retrieval company acquired by America Online Inc. The new service has close ties to the Internet Archive, a non-profit organization that is attempting to capture and maintain a digital snapshot of the World Wide Web.

The Alexa service attempts to monitor the behavior of Internet users systematically, the better to return information on the popularity of

particular Web pages. The idea, Mr. Kahle said, is to provide Internet users with some idea of the "context" of the page they view, as well as tips on where to look next.

This information emerges both from the preferences of other Alexa users and from analysis the service has done about the millions of links connecting different Web pages.

"We're trying to use the paths that people are leaving through the Net as an indicator of what is good," Mr. Kahle said.

Another advantage of the Alexa service is that it can retrieve "dead" Web pages that have been captured by the Internet Archive. The need for such capability has been growing as information often appears and then vanishes quickly.

Because the Alexa service, located at www.alexa.com, is free, it will be financed by advertisements that appear inside the program's display while the Internet user browses the Web.

Mr. Kahle said that since the service monitors individual user behavior on the Web, it has been designed to guarantee user anonymity in order to protect privacy.

That way, the more people who use *Alexa*, the better a navigation service it becomes, he said.

The idea for *Alexa* grew out of another of Kahle's projects, the Internet Archive, a effort to document and store Web pages and Usenet postings for a historical record and to provide, in the event of an outage, a sort of backup system.

In Thursday's outage of several of the Internet's name servers, for example, the Internet Archive existed as a sort of static mirror of nearly every Web page. The Archive currently has 5 terabytes of pages (five million megabytes).

Kahle is no stranger to developing new technologies for the Web. In 1989 he invented the WAIS technology for searching the Web, and founded WAIS Inc., an electronic publishing company that he later sold to America Online. Before that he helped found Thinking Machines, a maker of supercomputers.

Related Sites

Following are links to the external Web sites mentioned in this article. These sites are not part of The New York Times on the Web, and The Times has no control over their content or availability. When you have finished visiting any of these sites, you will be able to return to this page by clicking on your Web browser's "Back" button or icon until this page reappears.

- [Alexa Internet](#)
- [AltaVista](#)
- [Firefly](#)
- [The Internet Archive](#)

[Home](#) | [Sections](#) | [Contents](#) | [Search](#) | [Forums](#) | [Help](#)

Copyright 1997 The New York Times Company

Looking for a Second Commercial Hit, Brewster Kahle Launches a Navigation Service

By Margaret McGehey

ONE OF THE Internet's pioneers has developed a navigation service that uses collaborative filtering techniques to help users find relevant information on the Web.

Brewster Kahle, creator of one of the Internet's first text retrieval systems and now the founder of a startup called **Alexa Internet**, last week introduced a client-side Web navigation service that works with Windows 95-based browsers from Microsoft and Netscape.

Once the user picks a Web site to look at, the Alexa server displays a toolbar at the bottom of the page that features a pop-up screen recommending Web sites related to the

one currently being viewed. For example, if the user starts at the Prodigy travel site, Alexa might recommend going to the Compaq Traveler site, the Council Travel site, and other travel pages.

The recommendations on where to go are based primarily on analysis of the paths that other surfers took from that site. Those analyses are carried out on a bank of computers running in Alexa's offices in the Presidio in San Francisco.

"Alexa tries to infer what worked and what people liked based on where they went, which makes sense," said Jerry Michalski of the newsletter Release 1.0.

The Alexa software will be free

to users. Alexa's major source of revenue will be targeted advertising in the form of banner ads on the toolbar's pop-up screen.

For Kahle, who founded Alexa last year using money he made by selling his text retrieval firm, WALS Inc., to America Online in 1995, the challenge will be gaining widespread distribution of the Alexa software. "It's partnerships that will make or break us," Kahle said.

Analyst Patrick Kneue of Jupiter Communications added that the marketing costs associated with tabling Alexa could be substantial. Alexa has 25 employees and \$5 million in venture backing.

"They have to aggressively pro-

mote offline and buy banners online to push people to their site," Kneue said. "They can't expect advertisers to just flock to them."

At least one executive from the area of collaborative filtering suggested that Alexa could have a tough time ahead of it.

"People are beginning to get comfortable with finding sites on the Web, and Alexa is really looking to solve a problem that isn't fundamentally anyone's," said Mark Goldstein, the former CEO of Net Avenue, who was acquired by Findfly last year. "From my perspective, Alexa is going to be very challenged as a standalone product."

search AGENTS, SEARCH SERVICES



Alexa Internet CEO Brewster Kahle

LINKS FROM THIS PAGE

Alexa Internet—www.alexa.com
International Standards Organization—www.iso.ch

JavaSoft—www.javasoft.com
Microsoft—www.microsoft.com

Search Infocache: www.webweek.com/search

Netscape—home.netscape.com
Sun Microsystems—www.sun.com

July 21, 1997

Web Week

5

July 21, 1997

Technology

BellSouth cuts off AirTouch in Brazilian wireless bid

Steve Ginsberg

AirTouch Communications is still looking for a Latin American toehold for its wireless services after BellSouth outbid the San Francisco-based company for South America's business capital, Sao Paulo.

At \$2.5 billion, BellSouth's winning bid was four times higher than the lowest acceptable bid set by the government. The bid was also four times higher than any wireless license for other South American markets.

AirTouch, which finished third in the bidding, is proceeding cautiously in South America. "We were not prepared to win at any cost," said Kathy Reinhart, director of investor relations. "We have strict investment hurdles."

AirTouch has bids in for five other Brazilian markets. It won't have BellSouth to contend with because of rules restricting license holders to only one major market in Brazil.

With its international push led by President Arun Sarin, AirTouch is dominant in Europe, and the cheap licenses it landed early on may have spoiled the company for an era of inflated license fees. In 1992, AirTouch got the license for Germany and didn't have to pay an auction price.

BellSouth is staking out Latin America after entering Argentina in 1989. It is already operating in eight countries, including Brazil.

Alexa Internet lands at the Presidio

Internet startups are usually confined to gritty South of Market Street walkups in San Francisco or viewless Peninsula offices. But the 20 employees at Alexa Internet have views that would make most seven-figure CEOs jealous. Alexa is among the first for-profit companies to land at the Presidio National Park in San Francisco.

Alexa, a spinoff of a nonprofit already housed at the former army base, focuses on web navigation. The company launched its web site this week. The company is the branchchild of Bruce Gilliat and Brewster Kahle.

EA's annual report is fun and games

Annual reports have long been not only dissertations on financials, but a way for companies to market themselves. Readers of Electronic Arts' fiscal 1997 annual report won't encounter a dollar sign until page 13. The report's cover states in stark black and white, "Warning: This annual report is neither interactive nor entertaining."

The next 10 pages work to dispel that notion, showing clips of dunking hoopsters, racing cars and attacking helicopters. A reference to an EA game on the page opposite the basketball clip states merely, "It doesn't let you dunk."

EA can afford to be jocular. The company's 1997 numbers were good. Revenues grew 17 percent to \$625 million, while net income increased to \$53 million--a 30 percent jump. EA gains came, in part, because it guessed right. Its alignment with Sony PlayStation accounted for 30 percent of net revenues, compared with just 9 percent in 1996. Net revenues from Sega dropped from 28 percent to 16 percent.

EA is putting its healthy cash flow to work. The company expects to complete its acquisition of underachieving Maxis for around \$125 million in stock by Sept. 1.

Oakland set to tap Oracle for contract

Oakland's city council was expected to select Oracle Corp. of Redwood Shores to solve the city's year 2000

software problems.

A city council committee selected Oracle over Informix for the \$16 million contract, the most lucrative computer contract ever awarded in the financially-strapped city. The software developer will address the city's year 2000 problem as well as connect its 1,700 desktop PCs so that departments can communicate with each other.

Several smaller Oakland tech companies will get a piece of the Oracle contract, including Lily Hu & Associates, F2 Technologies and Cordoba Corp.

Deal update: Vanstar, U.S. West make buys

- In its bid to offer more extensive technological services to the federal government, Vanstar is acquiring Virginia-based Systorex in a stock and cash deal. Pleasanton-based Vanstar will use 500,000 shares of stock and \$46 million in cash to buy its former partner, Systorex, a \$150 million company.
- U.S. West has become the single largest investor in San Mateo's Wire Networks, a content provider of women's issues on the Internet. The size of the investment was not disclosed. Wire will create content in several cities for U.S. West's DiveIn web service. Wire has raised \$8 million since its 1992 founding and went online in 1995 with Women's Wire. The company will grow from 45 to around 75 employees this year. Recent hires included chief financial and technical officers.

© 1997, *San Francisco Business Times*

[More News Columns](#)

7/21/97

RETURN TO TECHSERVER: NORMAL || LOW-GRAPHICS[Click Here](#)

New service tracks Web use

Copyright © 1997 Nando.net

Copyright © 1997 N.Y. Times News Service

SAN FRANCISCO (July 21, 1997 04:39 a.m. EDT) -- As a second generation of Internet search tools emerges, a company called Alexa Internet has started a service offering users more information about the popularity and location of the pages they view.

The service, which works with Internet browsers offered by Netscape Communications Corp. and Microsoft Corp., focuses on the relatively new idea of assisting users in navigating the World Wide Web.

The Alexa service was founded by Brewster Kahle, who was one of the founders of Thinking Machines Corp., and by Wais Inc., an information retrieval company acquired by America Online Inc. The new service has close ties to the Internet Archive, a nonprofit organization that is attempting to capture and maintain a digital snapshot of the World Wide Web.

The Alexa service attempts to monitor the behavior of Internet users systematically, the better to return information on the popularity of particular Web pages. The idea, Kahle said, is to provide Internet users with some idea of the "context" of the page they view, as well as tips on where to look next.

This information emerges both from the preferences of other Alexa users and from analysis the service has done about the millions of links connecting different Web pages.

"We're trying to use the paths that people are leaving through the Net as an indicator of what is good," Kahle said.

Another advantage of the Alexa service is that it can retrieve "dead" Web pages that have been captured by the Internet Archive. The need for such capability has been growing as information often appears and then vanishes quickly.

Because the Alexa service, located at www.alexacom.com, is free, it will be financed by advertisements that appear inside the program's display while the Internet user browses the Web.

Kahle said that since the service monitors individual user behavior on the Web, it has been designed to guarantee user anonymity in order to protect privacy.

[[Global](#) | [Stateside](#) | [Sports](#) | [Politics](#) | [Opinions](#) | [Business](#) | [Techserver](#) | [Health & Science](#) | [Entertainment](#) | [Weather](#) | [Baseball](#) | [Basketball](#) | [Football](#) | [Hockey](#) | [Sport Server](#) | [MADN](#)]

Copyright © 1997 Nando.net

Do you have some feedback for the Nando Times staff?



CLICK HERE
TO SOLVE
THE PUZZLE



NEWSWATCH

Tuesday, July 22, 1997

TRENDS ONLINE

Contents

Today's News and Analysis

- **Real Standards for Streaming Media**
Microsoft and Progressive Networks agree to partnerships.
- **Mail: Killer Java App?**
Lotus and Sun prepare Java mail clients for cross-platform corporate use.
- **Needed: Internet Emergency Plan**
A nationwide computer glitch underscores the Net's vulnerability.
- **Life After 50: Cyberspace-Style**
Baby boomers (and their parents) take to the Web in record numbers.

NewsWatch

Trends Online
Internet Trends
Industry Trends
Hardware Trends
Software Trends
Trends Index

New Products

NEWS.SCAN

Today's Headlines from the Mainstream Press

Intel has settled a class action suit brought in the wake of the disclosure last year that an error in testing led the company to overstate the speed of its Pentium microprocessors by about 10 percent. Intel denied that it disseminated false or misleading information, but agreed to provide a \$50 rebate on the purchase of Intel OverDrive processors to customers who purchased a PC containing a 120-MHz or 133-MHz Pentium chip between October 23, 1995, and January 5, 1996. [NYT 7/21/97 p. D6]

The Internet address pool is dangerously low. In the early 1980s the Internet was crafted with a system of 4.3 billion potential network addresses—roughly one address for every person on Earth. Yet today, with a mere 50 million users online, the network address pool is running dry. Without new addresses, the network's expansion will soon cease. Scientists have a solution in the wings, but as changes are put in place over the next two to five years, the Net is likely to be fragile. [SUMN 7/21/97 p. E1]

LinkExchange, a San Francisco start-up, gives small-time Web sites wider exposure by creating a marketplace in which they can swap advertising space on their own sites for space on others. It's one of the few examples of an Internet company with a business model based on how the Internet works instead of imposing an old-line model on cyberspace. [WSJ 7/21/97 p. B7C]

Alexa Internet is a new Web service that attempts to monitor the behavior of Internet users and reveal the popularity of particular Web pages. Started by Brewster Kahle, a founder of Thinking Machines, and by Wais, an information retrieval company acquired by America Online, the service uses the

NEW PRODUCTS

[PCs]

PC Labs gets its hands on AMD's K5 processor.

[SOFTWARE]

Six software packages for making or filling out electronic forms.

Read about the latest 3-D home improvement software from Books That Work and Softedge. [PERIPHERALS]

Four new DAT drives with mass storage capabilities.

Protect your workstation or server with a UPS.

Hewlett-Packard's multifunction OfficeJet 570 may be the last printer you'll ever need.

[INTERNET]

Use the Web to get to know your customers.

Build communities with Creative Technologies's Creative Inspire.

preferences of other Alexa users and analyses of the millions of links connecting different Web pages. [NYT 7/21/97 p.D4]

Excite is expected to announce today Mailexcite, a new Web-based electronic messaging service that gives users a free e-mail address and account. Users will be able to check their e-mail from any computer using a Web browser. [NYT 7/21/97 p.D9]

CNN/Sports Illustrated yesterday announced SportSim a free fantasy sports Web site where players can pit their dream teams against teams concocted by other players. [NYT 7/21/97 p.D10]

The bigger the Internet gets, the more new "communities" seem to pop up. Thousands of sites from CBS News to Yahoo sponsor chats every day. Others: PeopleLink, Talk City, Webchat Broadcasting System, and The Palace. [NYT 7/21/97 p.D6]

Louis Rossetto, a founder of Wired magazine, will step down as CEO of Wired Ventures, but will continue as editor-in-chief and publisher of Wired and chairman of Wired Ventures. [NYT 7/21/97 p.D9]

High-tech dairy and cattle farms use electronic monitors and notebook computers to track animals' weights, heights, food consumption, health, milk output, and beef quality. [NYT 7/21/97 p.D4]

Burning Man, a five-day celebration in Nevada's Black Rock Desert, is the holiday of the digerati. [NYT 7/21/97 p.D4]

Compiled by Carol Levin and Leslie Sonich

Previous Headlines

News Sources	
BW	Business Week
FRB	Forbes
FRT	Fortune
NYT	New York Times
SFC	San Francisco Chronicle
SJMN	San Jose Mercury News
USA	USA Today
WSJ	Wall Street Journal

[NETWORKS]

Fibre Channel connectivity can help eliminate network bottlenecks.

New Gigabit Ethernet products for ultrafast networks.



Search



Inter@ctive Week

July 22, 1997

FRONT PAGE

MAIN VIEW

[New Today](#)

INFRASTRUCTURE

[Top Stories](#)

[Carol Wilson](#)

CYBERMEDIA

[Top Stories](#)

NET TOOLS

[Top Stories](#)

[David Courcy](#)

BUSINESS

[Net Index](#)

[Net Strike Data](#)

[Randy Whitestone](#)

[Net Companies](#)

ANALYSIS

[Inter@ctive Week](#)

[Each Day](#)

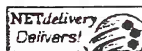
WEB GUIDE

RECRUITMENT

ZD NET TRADE SHOW CALENDAR

INDUSTRY EVENTS

SUBSCRIPTIONS



STHEAD

[Editorial](#)

[Sales](#)

[Operations](#)

New Guide Offers Net Directions

By [Tom Steinert-Threlkeld](#)

1:30 PM EDT

The next guide for directing Internet visitors to useful sites may be other Internet visitors.

The inventor of the pioneering Net publishing service known as the Wide Area Information Server system, Brewster Kahle, is at it again. This time, he is launching a service that tells Web users details about each site they visit and, at the same time, recommends other sites they may be interested in, based on analysis of surfing patterns of other visitors.

The new service, called Alexa, appears inside a toolbar that attaches itself to the bottom of a user's browser. "Screen real estate is certainly the most valuable real estate in the world," said Kahle. "We have to be worth more than real estate in Tokyo."

To achieve such worth, Kahle's Alexa Internet Corp., based in San Francisco, has been building up a huge archive of Internet pages. From the summer of 1996 to the present, Alexa has been sucking in any page that has appeared on the Web, building up a repository of 6 trillion bytes of Net information.

That will allow the company to track usage by Alexa-equipped browsers of not only current pages on the Web but also out-of-date pages. If a page can't be found at its original address, an Alexa-equipped browser can pull it up from Alexa's archive.


But to succeed, Alexa will need users. Its recommendations will be based on traffic patterns as captured by Alexa browsers. The toolbar will display sites of related interest, based on previous users' activities.

The bar also will provide pop-up access to data about the site being visited, such as who runs the site, where that person or company is located, how fresh the data is and other characteristics. That data will be linked to various other sources of information, such as maps and financial filings, that will allow users to better judge the quality of the site. Ratings of the sites also will be automatically generated, based on data collected by the Alexa browsers from users.

The browser add-on will be free, with registration to receive beta copies beginning last Thursday at the company's site. Alexa Internet will support itself through advertising, linked to various features of the toolbar.

Kahle's previous venture, WAIS Inc., was sold to America Online Inc. in 1995. It was based on Wide Area Information Server technology.

 can be reached at www.alexa.com

 Email [Tom Steinert-Threlkeld](#)

[Back to the top](#)

[Subscriptions](#)

[Ad Index](#)

[ZD Ad Marketplace](#)

[ZD Net](#)

The San Diego Union-Tribune.

(Page 21)

New search tool tracks user's path through Internet

New York Times News Service

22-Jul-1997 Tuesday

As second-generation Internet search tools emerge, a company called Alexa Internet has started a service offering users more information about the popularity and location of the pages they view.

The service, which works with Internet browsers offered by Netscape Communications Corp. and Microsoft Corp., focuses on the relatively new idea of assisting users in navigating the World Wide Web.

The Alexa service attempts to monitor the behavior of Internet users systematically, the better to return information on the popularity of particular Web pages. The idea is to provide Internet users with some idea of the "context" of the page they view, as well as tips on where to look next.

This information emerges both from the preferences of other Alexa users and from analysis the service has done about the millions of links connecting different Web pages.

Another advantage of the Alexa service is that it can retrieve "dead" Web pages that have been captured by the Internet Archive. The need for such capability has been growing as information often appears and then vanishes quickly.

The Alexa service (<http://www.alexa.com>) is free. It will be financed by advertisements that appear inside the program's display.

Copyright Union-Tribune Publishing Co.

7/23/97

internet news.com

internet advertising report

Daily Bulletin The Day's Top Stories

Compiled by Beth Cox
Internet Advertising Report

[Click here](#) to receive the Internet Advertising Report by e-mail!

News Alert Joins ADSmart Network

[July 23, 1997] New York City-based News Alert, an information service for investors, said that it has joined ADSmart.net and will begin offering custom sponsorships and interstitial advertising, in addition to standard banner advertising.

News Alert claims to deliver more than 3.5 million monthly page views, and 68% of all News Alert viewers trade stocks at least once a month.

Advertisers can now sponsor My Portfolio, News Alert's personalized portfolio, as well as personal news, live links to online trading sites, Trading corner and co-branded content on News Alert partner sites, including PC Quote, Individual Investor Online, Hoovers Online and Zack's Investment Research, Motley Fool, Edgar On-line, Wise Wire, Planet Direct and Nequity. Rate card data was not disclosed.

ADSmart.net is an advertising sales and management network specializing in custom-designed sponsorship opportunities for interactive advertising. ADSmart is a wholly owned subsidiary of CMG Information Services.

Auto-By-Tel, AllApartments Sign With Classifieds2000

Santa Clara, CA-based Classifieds2000 Inc. announced strategic agreements to distribute the content of Auto-By-Tel and AllApartments through its classified ad network. Financial terms of the deal were not disclosed.

"AllApartments and Auto-By-Tel are category leaders in their respective fields on the Internet, and we're pleased to offer them an opportunity to reach even more motivated, targeted users through the Classifieds2000 distribution network," said Sani El-Fishawy, CEO of Classifieds2000.

Classifieds2000 distributes its database of classified advertising content, display advertising and products to more than 50 Web publishing partners, including: search sites Excite, Lycos, InfoSeek and WebCrawler; top ISPs AT&T WorldNet, Concentric Network Corp., MindSpring Enterprises Inc., Netcom On-Line Communications Services Inc. and WebTV Networks Inc.; and popular category sites including Car and Driver and Road and Track.

US Interactive Launches Campaign for Newspaper Network

New York City-based digital marketing agency US Interactive has joined with the newspaper consortium New Century Network (NCN) and the latter's NewsWorks Web site for the launch of NCN's initial consumer advertising campaign online.

US Interactive said it created live, Java banner ads for the campaign. Billings were not disclosed.

The NewsWorks campaign will run through the first two weeks of September on Excite, Yahoo!, Netscape, Infoseek and Lycos. Creative executions were developed in-house by NCN.

"This campaign will aggressively establish the NewsWorks brand in the minds of Web users," said Paul Kessinger, vice president of marketing for New Century Network. "Not only are we targeting known Web users, but we're doing it in the two arenas netizens understand most: on sites that they frequently visit and via the use of engaging Java banners and animated gif banners."

Medizine, Mediconsult.com Sign Marketing Deal

MediConsult.com in Boston said that it has signed an exclusive content and marketing agreement with MediZine, a marketing company that reaches consumers through an education-based bi-monthly guidebook series distributed on pharmacy counters.

MediConsult.com will be MediZine's exclusive Internet partner and will develop a MediZine branded area on the MediConsult.com virtual medical center Web site. Financial terms of the deal were not disclosed.

MediZine will include MediConsult.com in its editorial and marketing materials for patients that are currently distributed through more than 8,500 retail pharmacies in the United States.

The agreement will enable MediConsult.com and MediZine to leverage advertising on the Internet, and across MediZine's 1.25 million health-specific guidebooks, the companies said. "Combining both print and Web advertising opportunities will generate greater ROI for potential sponsors through one-package offering," the companies said in a statement.

MediConsult.com is a "virtual" medical center on the Internet that offers peer reviewed medical information on more than 50 chronic medical conditions. The company partners with health care providers such as pharmaceutical companies, HMOs, hospitals, non-profit groups, physician's organizations and charitable research groups. Revenues are derived from products and service sales, Internet marketing consulting, content development, content-based advertising and educational grants.

I/PRO Guarantees 10-Day Audit Delivery

San Francisco-based [Internet Profiles Corp. \(I/PRO\)](http://Internet Profiles Corp. (I/PRO)) said that the company will now offer a 10-day delivery guarantee to Nielsen I/PRO I/AUDIT customers, surpassing what it called the industry's current practice of delivering audit statements up to 60 days after the specified audit period.

Said I/PRO's president and CEO Bradley Rode: "We can give customers and Web advertisers confidence that reports will be there when they need them while still having the reliable data that they have come to expect from I/PRO."

I/PRO said it has produced more than 5,000 audit reports since I/AUDIT was launched in 1995.

I/PRO analyzes, correlates, and validates Web activity that enables marketers to understand their customers and site usage. Clients include AT&T WorldNet, CBS, *Chicago Tribune*, CMP Publication Inc.'s TechWeb, Infoseek, Individual Inc., Nando.net, Starwave, Yahoo! and *USA Today Online*.

New Ad-Backed Service to Monitor Context of Web Pages

San Francisco-based Alexa Internet has started a new, ad-backed service offering users more information about the popularity and location of the pages they view.

The company was founded by Brewster Kahle, who was one of the founders of Thinking Machines Corp., and by Wais Inc., an information retrieval company acquired by America Online Inc., according to *The New York Times*. The new service has close ties to the Internet Archive, a nonprofit organization that is attempting to capture and maintain a digital snapshot of Web.

Alexa, which requires a download, promises "where to go next" suggestions and behind the scenes facts about the site you are visiting. Alexa works in parallel with, and independently of, Web browsers and is displayed as a toolbar on the user's desktop.

The service attempts to monitor the behavior of Internet users systematically, to return information on the popularity of particular Web pages. The idea, Kahle told the *Times*, is to provide Internet users with some idea of the "context" of the page they view, as well as tips on where to look next.

"We're trying to use the paths that people are leaving through the Net as an indicator of what is good," Kahle was quoted as saying.

Another advantage of the Alexa service is said to be that it can retrieve "dead" Web pages that have been captured by the Internet Archive.

Cable Channel Site Features Old TV Commercials

Banner ads are one thing, but will Web surfers tune in to see clips of old TV commercials? MTV Networks Inc. thinks so.

Video clips of classic TV commercials (can you say "Where's the beef?") will be featured on a newly launched combined Web site for cable channels Nick at Nite and TV Land, according to *Cable World* magazine.

Also on tap are program listings, descriptions and other promotional information for both Nick at Nite and TV Land. The site is running a consumer sweepstakes to draw computer users to its four electronic arcade games.

Internet Advertising Report Archives

Get the Internet Advertising Report by email once per week!

Enter your email address, then click SUBSCRIBE



301 Howard Street
Suite 1440
San Francisco, CA 94105
415.396.1800
415.396.1094 fax
info@antennapr.com



OUTSIDE THE BOX: Once upon a time, mainframes ruled the computing world and a group of men and women at a Cambridge, Mass., outfit called Thinking Machines decided to make a big, fast, fat computer that would function the way people work. Just last week a few of them thought they might be seeing that dream finally come true.

"It was an idea 10 years in the making," said **Lew Tucker**, who now manages software vendor relations at JavaSoft. Tucker was one of many Thinking Machine alumni at a party held to celebrate the launch of a new service, **Alexa Internet**, started by Thinking Machines founder **Brewster Kahle**. "The saying was, 'We wanted to build a machine that would be proud of us,'" he joked.

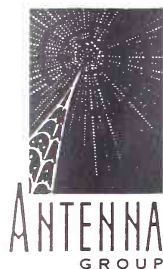
That idea, of course, failed. But some of the theorizing that Thinking Machines encouraged -- stuff about networks and artificial intelligence -- has helped make the Internet and the World Wide Web what it is today. And while the grand idea has mutated, the computer scientists' joy in being partly right is undiluted.

The geeks were so happy they were dancing at this reunion, cocktail party, product promotion and demo rolled into one happy event at the well-lighted and militarily sparse officers' club at The Presidio of San Francisco, near Alexa's offices.

Alexa is an Internet and Web search tool that's trying to be as smart as the people using it. Instead of letting users blindly guess the links between Web sites, Alexa's software refines a search, taking information about where other users have gone to point out areas of particular interest for those on the same path.

"This is Thinking Machines crossed with the Internet and the Web. We have enough users and enough data to make something happen," Kahle said. "We can live to replicate the Thinking Machines ideal."

Posted at 3:59 p.m. PDT Wednesday, July 23, 1997



301 Howard Street
Suite 1420
San Francisco, CA 94103
415. 896. 1800
415. 896. 1094 fax
info@antennapr.com

2 Internet Giants Won't Show at Expo

Microsoft, Netscape are skipping Chicago

Chicago

If the Internet industry puts on a conference and Microsoft and Netscape Communications don't bother to show, will the event make more than a peep?

The answer is yes and no.

Organizers of Summer Internet World 97, which starts today, insist the dynamic duo of the Net set will not be missed. They expect 25,000 people and more than 300 companies to attend the show, which is being held in Chicago for the first time. Internet World also is presented in New York in the winter and Los Angeles in the spring.

"It will take about a year to establish this show in the Midwest," said a spokesman for Mecklermedia, the Westport, Conn., publisher that puts on Internet World. "We expect the attendance figures to approach the New York and L.A. shows (50,000 each) next year. Remember, we had to move the show from San Jose to L.A. to handle the big crowds."

Seven months ago, Microsoft and Netscape were competing for headlines at Internet World in New York. And both also exhibited at the spring show in Los Angeles. This time, however, they're staying home, having concluded that Chicago amounts to overkill.

Officials of both companies said there are too many Internet-related shows these days. A spokeswoman for Netscape said the Mountain View company would rather use other resources — its online site and press releases, for example — to get its message out.

Without two of the Net's biggest players, the show has the feel of Chicago without Bulls merchandise on every street corner.

Meanwhile, other companies are eager to fill the vacuum and grab some attention.

■ IBM will introduce an "Emergency Response Team" of experts to fend off on-

ON TECHNOLOGY

Jon Swartz

line security attacks on enterprise networks. Big Blue will also offer a Web-hosting program for smaller companies and demonstrate something called the San Francisco Project, a Java programming initiative.

■ Oracle, the big database company, is adding payment options to its Internet commerce platform, making it easier for companies to buy software and services online.

■ Sun Microsystems will showcase the latest in Java, its programming language for the Internet. It also will demonstrate products designed for electronic commerce.

■ Apple Computer is introducing a Java version of its popular QuickTime video-editing program.

■ World Access is taking the lens cap off KidCam, an Internet video-conferencing system that allows parents to monitor the daily doings of their children at day-care and preschools.

■ Alexa Internet is announcing a free navigation service that gives users advice on where to surf the Web for specific information.

■ WhoWhere?, which maintains an online directory of information on people and businesses, is showing a free e-mail service called MailCity.

■ MapInfo is hawking an Internet mapping program that allows companies to display the geographic locations of their stores and offices to Net users.

The Oakland Press
Pontiac, MI
July 23, 1997

NEWS BYTE

Service turns page in aiding Web users

THE NEW YORK TIMES

SAN FRANCISCO — A second generation of Internet search tools is growing.

Alexa Internet offers a new service to provide more information on the popularity and location of World Wide Web pages.

The service works with Internet browsers from Netscape Communications Corp. and Microsoft Corp.

The Alexa service has close ties to the Internet Archive, a nonprofit organization that wants to maintain a digital snapshot of the Web.

The new service seeks to monitor the behavior of Internet users systematically and assess the popularity of particular Web pages.

The idea is to inform Internet users about the "context" of the page they view, as well as tips on where to look next, said Brewster Kahle, a founder of Thinking Machines Corp.

He started the Alexa service with Wais Inc., an information retrieval company acquired by America Online.

Alexa gleans its information by studying users' preferences and analyzing the millions of links connecting different Web pages.

"We're trying to use the paths that people are leaving through the Net as an indicator of what is good," Kahle said.

The Alexa service also retrieves "dead" Web pages that have been captured by the Internet Archive. Such a need has been growing as information often appears and then vanishes.

The service — located at www.alexa.com — is free. It will be financed by advertisements appearing inside the program's display while the Internet user browses the Web.

Since the service monitors individual user behavior on the Web, Kahle said, it has been designed to guarantee user anonymity and protect privacy.

Interactive Services Report
Washington, DC
July 25, 1997

- Getting even more personal: Her name is Alexa, and she is a new free Web navigation service from San Francisco-based Alexa Internet that learns about individuals' likes and dislikes and helps users determine whether or not a site is worth visiting. Alexa, which is now in beta, also offers suggestions about where a user might find the desired information based on an analysis of usage patterns, provides content and site links, and locates pages that no longer are available on the Web.



✕ As Go Surfers, So Goes Alexa

by Chris Oakes

5:03am - 24 Jul 97 PDT The Web wants to tell you a little bit about itself. At least that's the mission of a new service from San Francisco start-up [Alexa Internet](#). A free, advertising-supported "intelligent navigation service," Alexa acts as a browsing companion, providing contextual site information via a slim Windows toolbar.

The toolbar software, available in a beta version from Alexa's [Web site](#), receives and presents information independently from a user's browser and offers a range of basic data and site ratings.

"It's a navigation service that gets better as the Web grows and as there are more users," said Alexa Internet president and co-founder Brewster Kahle, who also founded WAIS Inc., an electronic publishing company sold to America Online in 1995. "Other people have found good things and not so good things," he said. "And as the discernment of users gets better, Alexa gets better."

The "meta-data" provided by Alexa is oriented around two questions: Where am I and where should I go next? The first question is answered via such details as ratings of a site's server speed, popularity, and content freshness, and what paths users tend to follow through the site. The second question is answered with a list of links to related sites others have frequented.

In drawing conclusions about sites, Alexa bases its ratings and recommendations on an analysis of patterns within the Web at large - as indicated by the preponderance of links to a particular site, for example. It also builds its opinions by tracking the usage "paths" of other Alexa users.

Jerry Michalski, managing editor of industry newsletter *Release 1.0*, likes what he sees and says Alexa has no direct parallel, even in agent-oriented sites recommendation services like [Firefly](#).

The general browsing population, Michalski believes, will take to it when it sees that rather than an overview map provided by directories like Yahoo, Alexa is a map with a localized structure. "Wherever you happen to be, it gives you the major roads out," he said. "That's pretty useful."

Alexa's Web analysis is based on a massive "copy of the Web" - an archive of all sites that have existed since early 1996. Already exceeding 5 terabytes in size, the company says the archive is updated with a new Web "snapshot" about every 60 days.

But Alexa's techniques raise some questions about the service's mechanics and even user privacy. For one thing, it's a service that gains intelligence as it is used, and its ratings can't yet draw on a wealth of Alexa-tracked "usage paths." To remedy this, Alexa editors are manually ensuring that "top Web sites" start out with appropriate ratings.

As far as the privacy of information on users' paths through the Web, Kahle says Alexa doesn't know who the users are - only what they're doing. "It's not that we don't exchange information [on a user's identity and behavior] - we don't have the information to exchange."

An additional feature similar to Excite's PAL chat service or

Click here now
InternetPaper.

PSINet
Better Net Results.

[Playboy Renews Fight
against Telecom Act](#)



TECHNOLOGY
Today's Headlines

[As Go Surfers, So Goes
Alexa](#)

[Java Pours from Many
Pots](#)

[Computers Steer
Automated Highway
Test](#)

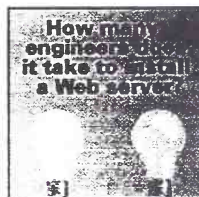
[Broken Glass, Sharp
Temper](#)

[Mac OS 8 Arrives](#)

[The Internet Ate My
Husband!](#)

[Scans: Most Valuable
Player](#)

[Tools: Internet Explorer
4.0 Preview 2](#)



AOL's instant messaging, Alexa's chat system can tell a browser who's online, and let them send messages in real-time.

Forrester Research senior analyst Mark Hardie says Alexa's success depends on its positioning versus directories and search engines. Alexa will have to make it clear that it's a navigation hub, educating Web users that "we're not about searching, we're about helping you get around."

The Web, as it continues to explode in content, may require the intelligence of a mass surfing population to help assess its content, as directories can only tally so much. "The richest directory that we know of points to less than 1 percent of all Web pages," Kahle said. That, he believes, leaves a lot of room for his service's approach.

Related Wired Links: -----

Fish4It Finds Tires, Pearls
9.Jul.97

Jeeves Finds the Way
9.Jul.97

Agents Are Getting to Know All about You
9.Jul.97



Feedback: Let us know how we're doing.

Tips: Have a story or tip for Wired News? Send it.

Copyright © 1993-97 Wired Ventures, Inc. and affiliated companies.
All rights reserved.

 **DELIVERED**

 **HOTWIRED**

COMPUTERS & TECHNOLOGY

**Alexa
Internet
hopes to
change
the way
you use
the Net**

By Zachary Coile
OF THE EXAMINER STAFF

Building 116 on the south end of the Presidio is a strange place to start a revolution. The two-story house with a hip roof and shipslapsiding was once a post trader's store and later noncommissioned officers' quarters.

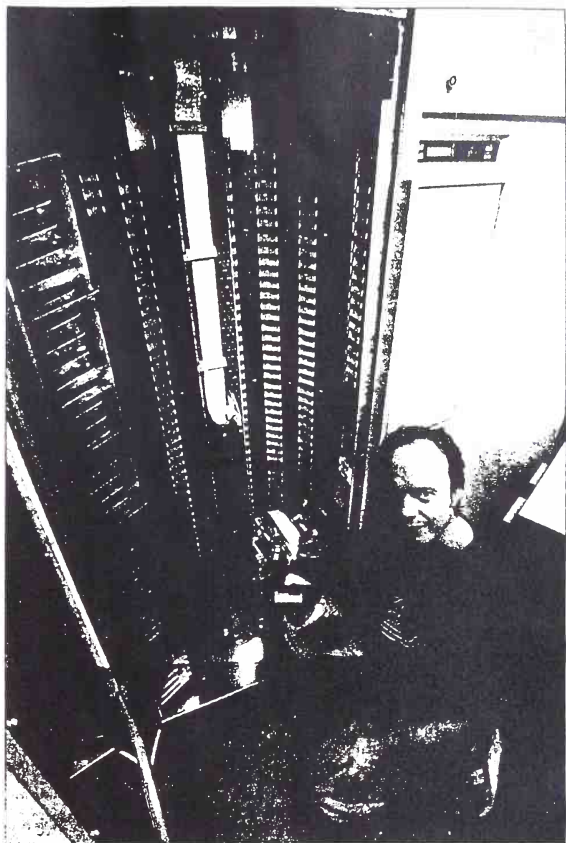
Alexa Internet, a small Web navigation company, moved into the building last year after the Army handed the Presidio over to the National Park Service. Now the building, built circa 1885, is ground zero for the company's new Web search service, which its founders say could change the way people use the Internet.

"The Web made everyone a publisher," says the company's founder and president Brewster Kahle. "Alexa allows everyone to be an editor." Or as co-founder and general manager Bruce Gilliat boasts: "We're the first and only Web navigation service that actually came from the people who use it."

One would normally dismiss that sort of talk as bluster, especially from a company with no revenues and just two dozen employees. But Kahle, a data mining specialist, and Gilliat have a record of making a success out of companies no one expects will survive.

Alexa's service appears as a thin, gray toolbar that can be moved around the screen while you browse the Web. The service is free to users and will be advertised supported. Currently being tested on those whose download of the Web, it will be officially released by the end of summer.

The company is positioning the service as an alternative to commercial directories and popular search engines like Yahoo and Excite.



EXAMINER/JOHN STOREY

Z Smith, vice president of engineering for Alexa Internet, kneels before the 20-terabyte digital tape machine in the company's temperature-controlled systems room. According to Smith, the machine can store all the text in the Library of Congress, with room to spare.

alternative they claim will make it easier for users to find what they're looking for.

What differentiates Alexa's service is the way it finds the information. Instead of using keyword searches, as most search engines do, or hiring workers to pick the best sites, as many directory services do, Alexa watches where Internet users go on the Web and uses the sum of their choices to make recommendations.

Bruce Gilliat demonstrates this capability by entering a site on surrealism. Clicking on a "Where to Go" icon, he pulls up a list of 10 related sites.

including Women Surrealist Artists, Dada, Alberto Giacometti and the Surrealists, and The Surrealism Test Center. All are sites that, users linked to from the original site.

"These are sites you may not have found with a search engine," Gilliat said.

Some of the seed money for Alexa Internet comes from Kahle and Gilliat's last venture, WAIS Inc., an electronic publishing company they sold to AOL for \$15 million in 1995.

The pair drew much attention nationwide last fall when they launched the Internet Archives, an ambitious attempt to record the entirety of the Web on a database for present and future scholars and researchers.

The vast archives are still stored on a 20-terabyte digital tape machine in Alexa's temperature-controlled systems room. The machine looks like a cross between a jukebox and vending machine, but it can store all the text in the Library of Congress and then some. The archives also have a commercial component for Alexa. By keeping snapshots of the Web that are updated every six weeks, the company can find pages that have disappeared, eliminating the dreaded "404 Not found" error. "Others

(See DATA, B7)

Revolution
brewing
at Presidio

◆ DATA from B-5

New revolution brewing at Presidio

are gathering the data," said Z Smith, vice president of engineering for Alexa. "The unique thing is that we can store the data and give it back to people."

Alexa's service also features what the company calls "contextual navigation." The idea is that users ought to know a little background on the site they're visiting.

To show what he means, Gilliat enters a site on a travel agency in New Jersey. Knowing nothing about the agency, he opens a box from the Alexa toolbar that gives some vital details: who owns the site, how many people have visited it, how frequently it is updated, how many Web pages there are, how fast the site's server is. The company also has a button where users can vote whether or not they liked the site. "There's no editorial

value to it," says Gilliat. "It's just what our users say."

Alexa users can also click on a box that lets them send messages to friends who are on-line, similar to services like ExcitePal, iChat's iPage and AOL's Buddy List.

Kahle said the company's service can be used along with other search services. "You still need some medium to get started, like a URL or a search engine," he said. But once people get to the site and start using Alexa's service "they may leave and not necessarily come back."

Joe Krause, senior vice president of marketing and co-founder of Excite, said Alexa may be the first to try such a service, but that all the major navigation services have the technology to offer users similar features.

Krause said that where Alexa might face an uphill battle is in distributing its service. "The key in our business is distribution. Getting eyeballs for your service is the

most important thing," he said. Excite has relationships with Microsoft, Netscape and America Online — the kind of partnerships Krause said Alexa may have trouble negotiating.

Still, Gilliat said advertisers have expressed interest in Alexa because it allows them to target users by their interest. "Mobil, for example, can say, 'Wherever there's automotive site, put our ad up there,'" he said. "The ads are content-sensitive to what you're looking at."

The company is allowing 10,000 people try the beta version of the service. "We have to get feedback and users and make sure our service works," said Gilliat.

And if the company is successful, will the founders sell again? "We did WAIS and sold it for \$15 million," he said. "We want to do things differently this time. We don't want to sell it right now. Maybe we'll have an IPO at a later time."

Plainview Herald
Plainview, TX
July 28, 1997

**BITS &
BYTES**



INTERNET CONNECTION
MODEM: 291-0000
TECH SUPPORT: 296-1354

**New service tracks
Web use**

SAN FRANCISCO — As a second generation of Internet search tools emerges, a company called Alexa Internet has started a service offering users more information about the popularity and location of the pages they view.

The service, which works with Internet browsers offered by Netscape Communications Corp. and Microsoft Corp., focuses on the relatively new idea of assisting users in navigating the World Wide Web.

The Alexa service attempts to monitor the behavior of Internet users systematically, the better to return information on the popularity of particular Web pages. The idea is to provide Internet users with some idea of the "context" of the page they view, as well as tips on where to look next.

This information emerges both from the preferences of other Alexa users and from analysis the service has done about the millions of links connecting different Web pages.

Because the Alexa service, located at www.alexa.com, is free, it will be financed by advertisements that appear inside the program's display while the Internet user browses the Web.

(JOHN MARKOFF/
N.Y. Times News Service)

Austin American-Statesman
Austin, TX
July 26, 1997

INTERNET WORLD '97

Net growth virtually exploding

■ The swelling digital bubble of Net and Web use in no danger of bursting, say enthusiasts

By JON SWARTZ
San Francisco Chronicle

CHICAGO — The Internet's meteoric rise is far from over. The party is just beginning.

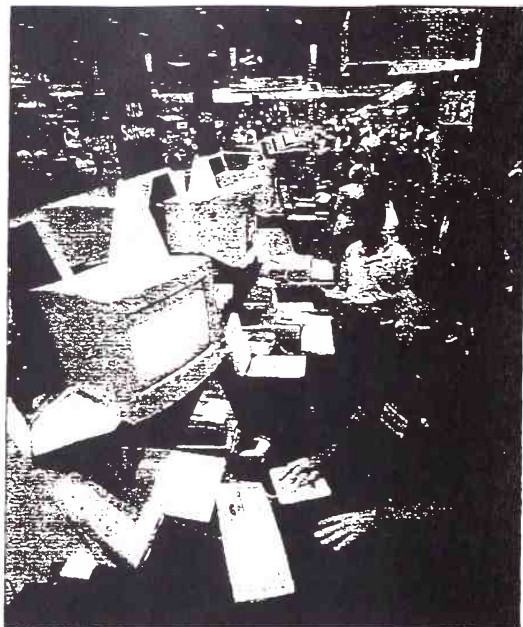
Those were the unfettered observations of revelers at the Summer Internet World '97 show here this week. They say the 2-year-old industry is not even close to peaking in popularity — in fact, it's on the cusp of another boom period.

"We see no end in sight," said Alan Taffel, vice president of marketing and business development at UUNET Technologies Inc., a Fairfax, Va., Internet service provider. "This is just the start of a long-term phenomenon."

The biggest opportunity, industry observers said, is in big business, where many corporations still rely on leased lines and public data services for communications. Many are just now turning to the Internet, after balking because of concerns over online security and quality of service.

Others said the integration of "killer apps" — software programs that improve video, sound and fax services on the Net — will fuel even greater growth in the corporate and consumer markets.

In a keynote speech on Wednesday, Novell Inc. chief executive Eric Schmidt said the Web will provide companies



Charles Bennett/AP

Rick Jackson of Digital Knowledge Assets works at a bank of computers at the Summer Internet World '97 on Wednesday in Chicago. E-commerce is the buzzword at this year's show.

with a "service model" for "information utilities."

All told, the momentum of the Net appears to be unstoppable, said showgoers at Internet World.

"The rate of change on the Internet is like dog years. For every year in the PC market it's like seven on the Internet," said Brewster Kahle, chief executive of San Francisco-based Alexa Internet, one of 360 exhibitors at the show. "It is a blistering pace."

UUNET has come up with its own phrase to describe the widespread acceptance of the Net: "Internet Law."

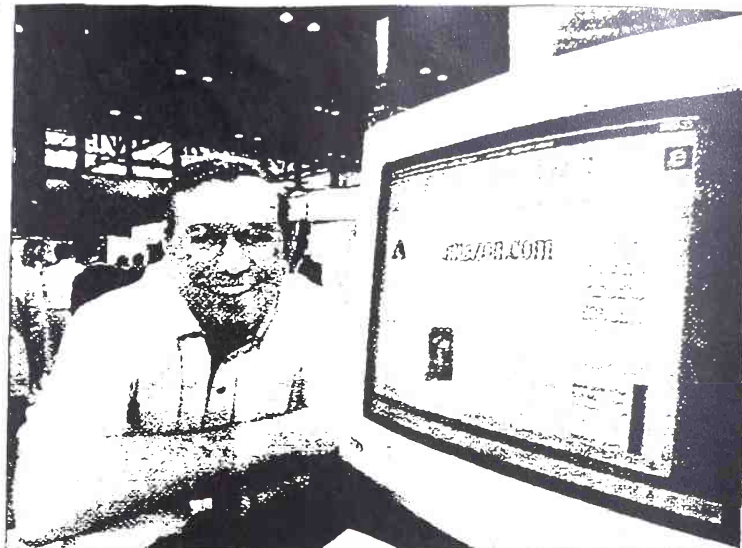
Based on Moore's Law, which states that the performance of a personal computer will double every 18 months at

the same price, "Internet Law" claims that the data-transmission capacity demands of consumers doubles every three to four months.

"Our network has to grow by a factor of 10 every year," Taffel said. "What other industry is even close to that rate?"

The Internet's rapid growth has forced industry heavyweights such as IBM Corp., Microsoft Corp. and Hewlett-Packard Co. to scramble to weave the Web into their sprawling business operations.

"The Internet evolution is creating a business revolution," said Dave Hansen, Internet business development manager at Palo Alto-based HP.



Associated Press

Charles Liberty, a marketing manager for Massachusetts-based Digital Equipment Corp., enjoys the first day of Internet World Summer '97, the industry trade show at Chicago's McCormick Place. "This year and next are probably the years for Internet commerce to take off," he said of an E-commerce boom that experts say may generate \$6 billion in revenue annually by the year 2000.

E-commerce about to launch?

By JON SWARTZ
San Francisco Chronicle

CHICAGO — If the Internet industry puts on a conference and Microsoft and Netscape Communications don't bother to show, will the event make more than a peep?

The answer is yes and no.

Organizers of Summer Internet World '97, that started Wednesday, insist the dynamic duo of the Net set will not be missed. They expect 25,000 people and more than 300 companies to attend the show, which is being held in Chicago for the first time. Internet World also is presented in New York in the winter and Los Angeles in the spring.

"It will take about a year to establish this show in the Midwest," said a spokesman for Mecklermedia, the Westport, Conn., publisher that puts on Internet World. "We expect the attendance figures to approach the New York and L.A. shows (50,000 each) next year. Remember, we had to move the show from San Jose to L.A. to handle the big crowds."

Seven months ago, Microsoft and Netscape were competing for headlines at Internet World in New York. And both also exhibited at the spring show in Los Angeles. This time, however, they're staying home, having concluded that Chicago amounts to overkill.

Officials of both companies said there are too many Internet-related shows these days. A spokeswoman for Netscape said the Mountain

View company would rather use other resources — its online site and press releases, for example — to get its message out.

Without two of the Net's biggest players, the show has the feel of Chicago without Bulls merchandise on every street corner.

Meanwhile, other companies are eager to fill the vacuum and grab some attention:

- IBM will introduce an "Emergency Response Team" of experts to fend off online security attacks on enterprise networks. Big Blue will also offer a Web-hosting program for smaller companies and demonstrate something called the San Francisco Project, a Java programming initiative.

- Oracle, the big database company, is adding payment options to its Internet commerce platform, making it easier for companies to buy software and services online.

- Sun Microsystems will showcase the latest in Java, its programming language for the Internet. It also will demonstrate products designed for electronic commerce.

- Apple Computer is introducing a Java version of its popular QuickTime video-editing program.

- World Access is taking the lens cap off KidCam, an Internet video-conferencing system that allows parents to monitor the daily doings of their children at day-care and preschools.

- Alexa Internet is announcing a free navigation service that gives users advice on where to surf the Web for specific information.

The Desert Sun
Palm Springs, CA
July 24, 1997

Beyond the valley

Internet industry giants snub expo

Online extravaganza:

With Microsoft and Netscape absent, other companies grab spotlight.

By JON SWARTZ

San Francisco Chronicle

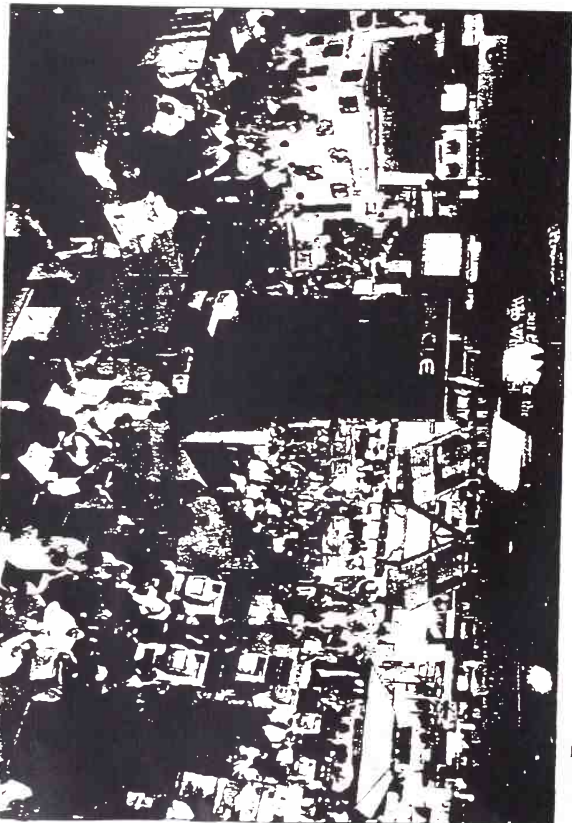
CHICAGO If the Internet industry puts on a conference and Microsoft and Netscape Corporation's don't bother to show, will the event make more than a prepp?

The answer is yes and no. Organizers of Summer Internet World '97, which started Wednesday, insist the dynamic duo of the Web are still not to be missed. They expect 25,000 people and more than 100 companies to attend the show, which is being held in Chicago for the first time. Internet World also is presented in New York in the winter and Los Angeles in the spring.

It will take about a year to establish this show in the Midwest," said a Spokesman for Microsoft, under the Webtop, CompuLink. We expect the attendance figures to approach the New York and L.A. shows for this year's event. He said they had to move the show from San Jose to L.A. to handle the big crowds.

Seven months ago, Microsoft and Netscape were competing for headlines at Internet World in New York. And both also exhibited at the spring show in Los Angeles. This time, however, they're staying home, having concluded that Chicago amounts to overkill.

Officials of both companies said there are too many Internet-related shows these days. A spokeswoman for Netscape said the Mountain View company would rather use other resources — its online site and press releases, for example — to get its message out.



Computer show: Companies display wares Wednesday at Summer Internet World '97 in Chicago.

IBM, Network security, selective Web surfing, and electronic supervision are hot topics.

Meanwhile, other companies are eager to fill the vacuum and grab some attention.

IBM will introduce an "Enterprise Response Team" of experts to fend off online security attacks on enterprise networks. Big Blue also will offer a Web-hosting program for smaller companies and demonstrate something called the San Francisco Project, a Java programming initiative.

Oracle, the big database company, is adding payment options to its Internet commerce.

making it easier for companies to buy software and services online.

■ **Sun Microsystems** will showcase the latest in Java, its programming language for the Internet. It also will demonstrate products designed for electronic commerce.

■ **Apple Computer** is introducing a Java version of its popular QuickTime video-editing program.

■ **World Access** is taking the lead cap off KidCam, an Internet video-conferencing system that allows parents to monitor their children at day-care and preschools.

■ **AltaVista** is announcing a free navigation service that gives users advice on where to surf the Web for specific information.

■ **WhoWhere?** which maintains an online directory of information on people and businesses, is showing a free e-mail service called MeltCity.

■ **MapInfo** is launching an Internet mapping program that allows companies to display the geographic locations of their stores and of their Web users.

The expo runs thru Aug. 1.

Post-Bulletin
Rochester, MN
July 28, 1997

Officials see no end in sight for the popularity of Internet

By Jon Swartz
San Francisco Chronicle

CHICAGO — The Internet's meteoric rise is far from over. The party is just beginning.

Those were the unfettered observations of revelers at the Summer Internet World 97 show here this week.

They say the 2-year-old industry is not even close to peaking in popularity — in fact, it's on the cusp of another boom period.

"We see no end in sight," said Alan Taffel, vice president of marketing and business development at Unet Technologies Inc., a Fairfax, Va., Internet service provider. "This is just the start of a long-term phenomenon."

The biggest opportunity, industry observers said, is in big business, where many corporations still rely on leased lines and public data services for communications.

Many are just now turning to the Internet, after balking because of concerns over online security and quality of service.

Others said the integration of "killer apps," — software programs that improve video, sound

and fax services on the Net — will fuel even greater growth in the corporate and consumer markets.

In a keynote speech, Novell Inc. chief executive Eric Schmidt said the Web will provide companies with a "service model" for "information utilities."

All told, the momentum of the Net appears to be unstoppable, said show goers at Internet World.

"The rate of change on the Internet is like dog years. For every year in the PC market, it's like seven on the Internet," said Brewster Kahle, chief executive of San Francisco-based Alexa Internet, one of 300 exhibitors at the show. "It is a blistering pace."

Unet has come up with its own phrase to describe the widespread acceptance of the Net: "Internet Law."

Based on "Moore's Law," which states that the performance of a personal computer will double every 18 months at the same price, "Internet Law" claims that the data-transmission capacity demands of consumers doubles every three to four months.

"Our network has to grow by a factor of 10 every year," Taffel said. "What other industry is even close to that rate?"

The Internet's rapid growth

has forced industry heavyweights such as IBM Corp., Microsoft Corp. and Hewlett-Packard Co. to scramble to weave the Web into their sprawling business operations.

"The Internet evolution is creating a business revolution," said Dave Hinman, Internet business development manager at Palo Alto-based HP.

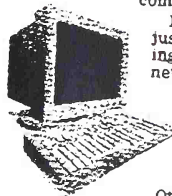
To harness the runaway growth, HP is in the process of Web-ifying its business model so that its printer, PC and server divisions are working in sync with its Net activities. The company is also stepping up efforts to market Internet products and services, Hinman said, although he

declined to say how.

Another indication of the growth of the Net is that spending on Web ads more than doubled in the first half of this year, to \$217.3 million, according to the newsletter Electronic Advertising & Marketplace Report.

Even the Internet trade-show industry seems to be thriving. Mecklermedia, the Westport, Conn., organizer of this week's show, announced plans to launch a Web conference for consumers next January in Boston.

It now runs four Internet shows a year — twice as many annually as Comdex and Macworld Expo.





301 Howard Street
Suite 1440
San Francisco, CA 94105
415. 996. 1800
415. 896. 1094 fax
info@antgroup.com

mediaCentral

July 28, 1997
Volume 3, No. 30

Media Central Digest

Alexa Internet -- One of the better tools we've seen for Web surfing lately is Alexa -- a browser aid now in beta. Alexa gives users a heads up on how many pages are on the site, who owns it, and a relative gauge of how popular a site is by how many pages out on the Web refer to it. In addition, if you go to a link that leads to a dead page, chances are Alexa may have recorded a copy of it. Finally, Alexa serves an agent function: Because it has access to a data mine of path information, it's able to make suggestions based on the "beaten paths" left by other surfers who've hit the same sites. Finally, users can rate sites themselves and communicate with other Alexa users. The company just started trickling out betas of its software last week, and so far we like it. It allows site seers to get their bearings on the Web in what struck us as a new and user-empowering way.

CYBERMEDIA

Online City • JULY 23, 1997 • 11



ANTENNA GROUP
301 Howard Street
Suite 1440
San Francisco, CA 94105
415.896.1800
415.896.1094 fax
info@antgroup.com

Alexa Keeps Surfers On Beaten Path

By Tom Steinert-Threlkeld

With 640,000 sites and counting, the Web remains a tangled knot of information for many visitors.

In the eyes of Brewster Kahle, the fastest way to get users where they're going may be to use a relatively untapped set of Net experts — other Web visitors.

Kahle helped start Thinking Machines Corp. — which figured out a way to tie together thousands of processors to create powerful computing engines — and invented the pioneering Net publishing service known as the Wide Area Information Server system. Now he has turned his attention to creating a new means of adding relevance and intelligence to navigation of the Web.

He is launching a service to offer Web users details about each site they visit and to recommend other sites, based on analysis of the surfing patterns of other visitors.

The new service, called Alexa, appears inside a tool bar that attaches itself to the bottom of a user's browser.

"Screen real estate is certainly the most valuable real estate in the world," Kahle said. "We have to be worth more than real estate in Tokyo."

To achieve such worth, Kahle's Alexa Internet Corp., based in San Francisco, has been building a huge archive of Internet pages. From the summer of 1996 to the present, Alexa has sucked in any page that has appeared on the Web, building a repository of 6 trillion bytes of information.

The cache will allow the compa-

ny to track the popularity of pages and to offer suggestions of other sites that users might like. The service will be available to Webmasters, an Alexa Internet Corp. spokeswoman said.

But Kahle says the service will need

and financial filings, that will allow users to better judge the quality of the site. Alexa will also automatically generate ratings of the sites based on data collected from users.

But does the Net need another rating service? Navigation services, such as those from Yahoo! Inc. and Excite Inc., provide increasingly powerful search and directory services. Almost any news service of any repute provides recommendations of related links. Pieces of software called agents are being refined to let users predetermine what information they want to retrieve and then collect it automatically.

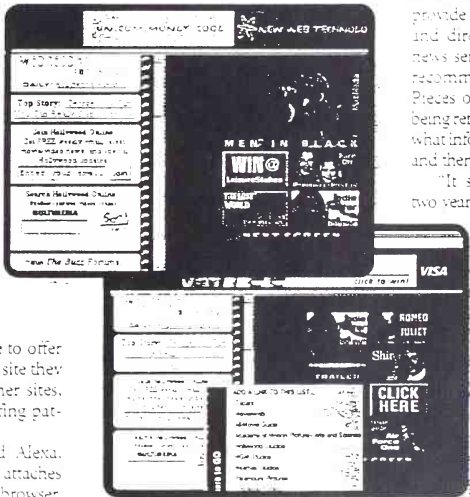
"It seems like something from two years ago," says Peter Krasilovsky, senior analyst at Arlen Communications Inc. in Bethesda, Md. "It seems like they're assuming Internet users need a very strong tool to help them find things."

The browser add-on will be free to the first 10,000 users who request it. Registration to receive beta copies began July 17 at the company's site (www.alexa.com).

Alexa Internet will support itself through ads linked to features of the tool bar. But Kahle says no advertisers have yet been signed.

Kahle founded the company along with Bill Dunn, a former online services executive at Dow Jones & Co. The company, started in April 1996, received \$5 million in funding from Etoile Investments of Switzerland.

Kahle sold his previous venture, WAIS Inc., to America Online Inc. in 1995. ▲



WEDNESDAY, JULY 23, 1997



ANTENNA GROUP
301 Howard Street
Suite 1410
San Francisco, CA 94103
415 896 1800
415 896 1094 fax
info@antgroup.com

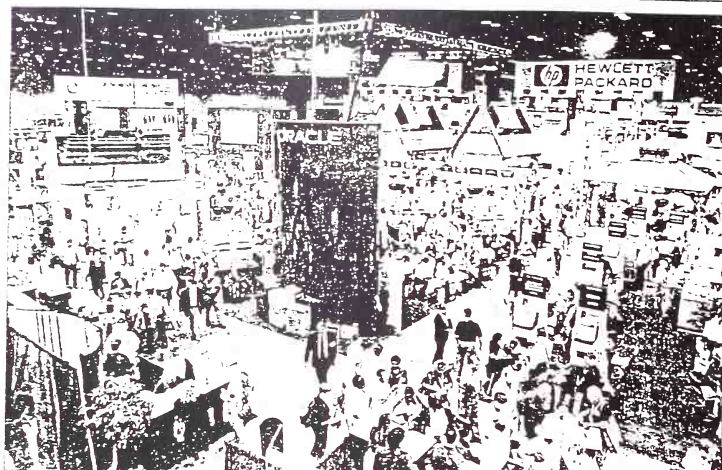


PHOTO BY ASSOCIATED PRESS

Industry revelers schmaazed and wandered about the various stands at Summer Internet World 97 in Chicago.

The Internet Boom Is Just Beginning

Big business is the next frontier, attendees at Chicago trade show say

By Jon Swartz
Chronicle Staff Writer

Chicago

The Internet's meteoric rise is far from over. The party is just beginning.

Those were the unfeigned observations of revelers at the Summer Internet World '97 show in Chicago last week. They say the 20-year-old industry is not only poised to peak in popularity, but also to launch the next boom in computing.

"We're not out of sight," said Alan T. Moore, president of marketing and business development at United Technical Resources, a Fairfax Va. Internet services company. "This is just the start of a long-term phenomenon."

Those 200,000 opportunity industries of servers said, in big business, where many corporations still rely on leased lines and proprietary data services for communications. Many are just now turning their backs on after talking the talk of e-commerce, file security and quality of service.

Others are in the transition of making their own infrastructure that im-

proves their sound and fax services, the Net will fuel even greater growth in the corporate and consumer markets.

In a keynote speech on Wednesday, Novell Inc. chief executive Eric Schmidt said the Web will provide companies with a "source model" for information utilities.

All told, the momentum of the Web appears to be unstoppable, said show organizers at Internet World.

"The rate of change on the Internet will be a surprise. For every year in the PC market, it takes six months on the Internet," said Frederick Keller, chief executive of San Francisco-based Alex Internet, one of 400 exhibitors at the show. "It is a blistering pace."

Internet has come up with its own phrase to describe the widespread acceptance of the Net: "Internet Law."

Robert Moore's Law, which states that the performance of a product doubles every 18 months, is a 1980s relic at this point, says Internet Law. "Items that the law forecasted doubling in performance would take three to six months."

Our test took less than a year to grow by a factor



Rick Jackson of Digital Knowledge manned a computer at the show.

of its size, Taffel said. "What other growth is even close to that rate?"

The Internet's rapid growth has led to a cluster of upstarts such as ITM, the Mountain View, Calif. and Hewlett-Packard, two giants to watch the Web for their sprawling business operations.

The Internet revolution is creating a new type of computer, says Dave Hinman, director of business development manager at P.H.I. based HP.

It is a new type of growth HP has seen. The Web, he said, is business's most important driver. PC and

INTERNET



July 28, 1997 10:00 AM ET

Alexa searches while you surf

By Jim Karstetter, PC Week Online

A new San Francisco company is putting a fresh spin on search technology: Search as you go.

Alexa Internet Inc. last week launched beta trials of a free service that, instead of giving users a static search results page, suggests places to go as users make their way around the Web.

The service provides background information about a site a user is visiting and suggests where a user may want to go based on an analysis of Web surfing patterns, content and linkages.

Alexa developers, some of them former employees of parallel processing machine maker Thinking Machines Corp., have spent the last year essentially cataloging the Web, drawing comparisons between sites and determining where users are likely to go while moving around.

Users navigate the search results with the Alexa Toolbar, which is available for download at www.alexa.com for Web browsers that run on Windows 95 or Windows NT.

The service draws information for four basic features: background information on the current site, Web surfing suggestions, a cached collection of dead Web sites stored on Alexa's Solaris servers and chat groups.

The beta period, which started last week, is limited to the first 10,000 users to register and download the client.

Alexa is relying on advertising for its revenue and should take the service to general release early next year.

"This is a pretty nifty idea," said Dan Chen, an IT manager at a San Francisco company. "I hope they come up with enough advertising to keep it going."



TOP OF PAGE

— A ZDNet Site —

[ZDNET HOME](#)

[SITE GUIDE](#)

[FIND IT](#)

[COMMUNITY](#)

[AD INFO](#)

[CONTACT US](#)

Copyright 1997 Ziff-Davis Inc. All Rights Reserved. Reproduction in whole or in part in any form or medium without express written permission of Ziff-Davis Inc. is prohibited. PC Week and PC Week Online and the respective logos are trademarks of Ziff-Davis Inc.

[Send e-mail to PC Week](#)

PCWEEK

JULY 28, 1997 • Volume 14, Number 32

BREAKING STORIES: www.pcweek.com

NEWS

Alexa Searches While You Surf

BY JIM KERSTETTER

A NEW SAN FRANCISCO COMPANY IS PUTTING a fresh spin on search technology: Search as you go.

Alexa Internet Inc. last week launched beta trials of a free service that, instead of giving users a static search results page, suggests places to go as users make their way around the Web.

The service provides background information about a site a user is visiting and suggests where a user may want to go based on an analysis of Web surfing patterns, content and linkages.

Alexa developers, some of them former employees of parallel processing machine maker Thinking Machines Corp., have spent the last year essentially cataloging the Web, drawing comparisons between sites and determining where users are likely to go while moving around.

Users navigate the search results with the Alexa Toolbar, which is available for download at www.alexa.com for Web browsers that run on Windows 95 or Windows NT.

The service draws information for four basic features: background information on the current site, Web surfing suggestions, a cached collection of dead Web sites stored on Alexa's Solaris servers and chat groups.

The beta period, which started last week, is limited to the first 10,000 users to register and download the client.

Alexa is relying on advertising for its revenue and should take the service to general release early next year.

"This is a pretty nifty idea," said Dan Chen, an IT manager at a San Francisco company. "I hope they come up with enough advertising to keep it going." ♦



"All the Bits
That Do Not Byte"

Internet Outlook

International Edition
Today: Electron Storms
and a Million Bit Streams

EXTRA!!!

July 28, 1997

[home](#) / [experts](#) / [internet](#)

EXTRA! Highlights from Internet World Summer Show

By Richard Wiggins

Chicago hosted most of the major players in Internet technology last week when Mecklermedia brought its Internet World trade show to the Windy City. Mecklermedia hosts Internet World trade shows on the East Coast of the US in Fall (usually New York City) and the West Coast in Spring (usually Los Angeles). Last week's show was the inaugural summer Internet World, presumably to be held in Chicago each summer for a primarily midwest audience. The show was held in the sparkling new North Building of McCormick Place.

Internet Outlook is a biweekly column exploring Internet trends: what's happening with Internet technologies, and how individuals, businesses, and government are coping with changes that occur in "Internet time."

Microsoft Alliance Leads to Free RealAudio / RealVideo Servers

During Internet World, Progressive Networks announced that Microsoft will invest in the company and will license RealAudio and RealVideo technology for use in its Netshow product. Netshow made a splash this summer when Bill Gates gave the [keynote address at the Special Libraries Association meeting in Seattle](#). The alliance with Progressive means that Netshow will be based on the industry leader's technology instead of some new Microsoft concoction.

Paul Thelen, Progressive's marketing manager for RealVideo, told me that Microsoft also will install RealAudio and RealVideo server capability into its new Site Server releases. At the same time, Progressive announced that its streaming server will now be available for free for low-volume applications: any company can use one 60-stream server license without charge.

Finally, Microsoft will bundle RealVideo player software in all distributions of Internet Explorer -- including when the Explorer is pre-installed on PCs for retail sale. Thelen says some users complain about the overhead of having to install Real players, "and this means ubiquitous player distribution."

Progressive thus enhances its role as market leader on both the server side and the desktop. On the server side, any customer using Microsoft server software already has Progressive's technology ready to use. Any low volume site, which previously may have relied on the inferior alternative of pseudo-streaming, now serves for free. And on the desktop side, users can tap into Real formats without downloading and configuring a plugin. The net effect of Microsoft's moves solidifies Progressive's already formidable position as leader in the streaming multimedia arena.

A Rival to CU-SeeMe?

CU-SeeMe was written at Cornell University about five years ago by Tim Dorcey. Dorcey was working as a statistician at Cornell (he has a master's in statistics and a master's in psychology) and was challenged by his boss to write a program that took advantage of new Macintosh video capability, offering desktop videoconferencing via the Internet.

Dorcey accepted the challenge. Even though CU-SeeMe was his first serious programming effort, it became wildly popular. Eventually a Windows version was offered, and Cornell commercialized CU-SeeMe through White Pine Software.

Now, Dorcey comes forth with iVisit, marketed by a California firm, Boxtop. Dorcey says his second generation Internet video conferencing product is far better tuned to the Internet. In an interview he touted the fact that iVisit does not require reflectors for groups to set up a conference; he's developed a relay architecture that forwards audio and video bitstreams efficiently around a virtual circle of participants. This would make it easy for anyone to form a group of participants without having to install reflector software.

Dorcey also says he's made the problem of finding groups of people to communicate with "much more like the ways in which people meet people in real life." With CU-SeeMe, finding people is accomplished primarily through non-CU-SeeMe channels, such as Web pages, Usenet postings, and the Four11 service. If Dorcey's claims are true, a product that performs better than CU-SeeMe, that doesn't require reflectors, and solves the problem of finding people and affinity groups will prove formidable competition for CU-SeeMe.

Alexa: Merging Web Searching and Group Experiences

Also announced in time for Internet World was Alexa Internet, "the Web navigation services that learns from people." Alexa is the brainchild of Brewster Kahle, best known for having invented Wide Area Information Servers, or WAIS, when he worked for a supercomputer vendor, Thinking Machines Corporation.

Alexa is service that allows the Web surfing experiences of many users to be tracked centrally. The Alexa search engine takes into account how popular sites are when weighting results of searches to present to users. In order to use Alexa, you must download the Alexa client program, which works in conjunction with your Web browser. Kahle told me they experimented with Java for this function, but found they couldn't build an applet that performed as well as compiled client code.

The product was announced this month but client programs are being distributed on controlled release while the Alexa staff adds robustness to the service. Kahle demonstrated Alexa for me showing how individuals' surfing habits are added to the central Alexa database. Kahle emphasized that personal information is not gathered by the Alexa service: "We couldn't even honor a court order to reveal what your favorite sites are because we don't gather that data. We gather aggregate data, not individual data."

Alexa ties into another Kahle venture, the Internet Archive. This non-profit undertaking seeks to build an archive of all of the content on the Web (and on Usenet) on a regular basis. Kahle told me his motivation is "capturing our digital heritage for future generations." He says they now have several terabytes of data in the Archive. In general, the Archive is not open for mass exploration; however, a few scholars are already examining its contents.

Sometimes Web sites experience meltdowns of one sort or another. For instance, when the Heaven's Gate mass suicide became known, the cult's Web site was so overloaded that pages were inaccessible. In other cases, pages or entire sites may go offline permanently. Alexa will have access to the Internet Archive. Kahle says the Alexa user will not see the dreaded Error 404 for a page that's no longer accessible; instead, he or she will see a copy

from the Internet Archive.

The group surfing concept is a fascinating one. It's been discussed in various forms for as long as we've had a Web, but to my knowledge this is the first serious attempt at implementing the concept. I fear that there may be significant obstacles for Alexa to overcome: people don't want to install yet another client, and people will need a compelling reason to try yet another search engine. But if Alexa delivers on the promise of getting people the content they really want or need, its future is bright.

Digex: A Regional ISP Steps Up to the National Stage

Among Internet Service Providers at the show, Digex made the biggest splash. Anxious to shed their image as a D.C. area ISP, Digex offered an imposing booth, ran large ads in the show brochure, and held a press conference to tout their new strength after merging this month with Intermedia. Digex brags about 99.9% uptime, redundancy, and the fact that they only use Cisco routers. Obviously they feel that businesses will increasingly want ISPs that offer high reliability and good throughput, and they want business Internet connectivity buyers to think of Digex as leaders.

I chatted with Clyde Heitzelman, president of Digex' Business Internet Connectivity Group, about the role of ISPs in the future. Specifically I asked if telephone companies won't eventually wipe out a lot of ISPs, once they figure out how to offer good Internet services, simply because they own the phone lines. Heitzelman, whose worked many years for major telephone players, said that regulators will require phone companies to offer bandwidth to their Internet business units at the same prices external ISPs pay, and that external ISPs like Digex will be able to compete effectively.

Are Internet Trade Shows Changing?

The show seemed somewhat smaller than previous Internet Worlds. Both Netscape and Microsoft were conspicuous by their absence. It's not un-heard of for Netscape to skip a trade show, but usually Microsoft is there if it makes any sense to be there. Overall the number of exhibitors seemed somewhat less than at past events. Attendance seemed lower than one might have expected, and I noticed coupons in the Chicago Tribune offering free attendance. Perhaps as a new venue the Chicago show will need a year or two to ramp up.

The feel of this show was different than past Internet Worlds. Past shows would appeal equally to consumers, to Internet newbies, and to businesses. Here, electronic commerce seemed to be the biggest focus, with major technology firms such IBM, Sun, HP, and Oracle emphasizing tools for doing business over the Internet. Companies in the Internet business feel they will make money from other businesses, not from end consumers.

Turnstile counts at trade shows are as unreliable as police crowd estimates at riots, but Internet World claims to attract up to 50,000 attendees at each of its shows. With several shows in the US and around the world, more people may physically attend an Internet World show than the 300,000 who subscribe to the magazine of the same name. It seems neither a glossy magazine nor all the content on the Web can compete with the old fashioned face to face convention.



[Comments](#) are welcome

Produced by [Richard Wiggins](#) and Copyright © 1997 [Athenia Associates](#)

AUG.



AUG.

Alexa Launch Coverage: July and August, 1997

Headline:	Publication:	Date:	Author:
FYI: Popularity Context	Star Tribune (Twin Cities/Minnesota)	8/4	New York Times
Scent of a Web site	Inman News Features	8/4	Inman News
Net Surf: Alexa's New Navigation Service	Wired News	8/1	Joey Anuff
Smart Searching at Internet World (Web Addict Section)	Web Review	8/1	Wes Thomas
Site Reviews: Alexa Internet	Media Central Digest	7/29	Sean Wolfe
Alexa: Merging Web Searching and Group Experiences	Internet Outlook (http://www.webreference.com)	7/28	Richard Wiggins
Alexa Keeps Surfers on Beaten Path	Inter@ctive Week (print)	7/28	Tom Steinert-Threlkeld
Alexa Searches While You Surf	PC Week	7/28	Jim Kerstetter
Revolution Brewing at Presidio: Alexa Internet hopes to change the way you use the Net	San Francisco Examiner	7/27	Zachary Coile
The Internet Boom Is Just Beginning: Big business is the next frontier, attendees at Chicago trade show say	San Francisco Chronicle	7/25	Jon Swartz
Outside The Box	San Jose Mercury News	7/24	Chris Nolan
As Go Surfers, So Goes Alexa	Wired News	7/24	Chris Oakes

Headline:	Publication:	Date:	Author:
2 Internet Giants Won't Show at Expo: Microsoft, Netscape are skipping Chicago	San Francisco Chronicle	7/23	Jon Swartz
New Ad-Backed Service to Monitor Context of Web Pages	Internet Advertising Report	7/23	Compiled by Beth Cox
New Guide Offers Net Directions	Inter@ctive Week (online)	7/22	Tom Steinert-Threlkeld
New Search Tool Tracks User's Path Through Internet	The San Diego Union-Tribune	7/22	New York Times News Service
News.Scan: Today's Headlines from the Mainstream Press	PC Magazine Online	7/22	NewsWatch
Looking for a Second Commercial Hit, Brewster Kahle Launches a Navigation Service	Web Week	7/21	Margaret McKegney
Alexa Internet Lands at the Presidio	San Francisco Business Times	7/21	Steve Ginsberg
New Service Tracks Web Use	Nando.net (Reprint of NY Times Article)	7/21	John Markoff
New Service Tracks Web Use	New York Times	7/21	John Markoff
Alexa Internet: The Search as a Communal Effort	New York Times (CyberTimes Extra)	7/19	Laurie J. Flynn
Recent & Decent Web Tools	Family PC News Online	7/18	Michelle Megna
Alexa Browser Companion Makes Searches More Certain	PC World Online	7/18	Lisa Moskowitz
Re-engineering Surfing	San Jose Mercury News Online	7/18	Patricia Sullivan, Online Editor

Headline:	Publication:	Date:	Author:
	(link to TechWeb story)		
Internet Historian Creates Tool to Inform Web Site Visitors	ZDNN-- The ZDNet News Channel	7/18	Renee Deger
Alexa Debuts Search Engine That Learns from Its Users	ZD Internet MegaSite Magazine	7/18	no byline
Startup Seeks To Synthesize Web	TechWeb	7/17	John Gartner
Alexa Internet Introduces Web Navigation that Learns from People	LA Times, BizWire Section	7/17	Press Release
Alexa Makes the Web History	C/Net	7/9	Alex Lash
Personal Access: Information abounds on the Internet, which means virtually anyone can find out virtually anything about someone else	Tucson Citizen	6/30	Leslie Miller

BUSINESS ISSUES IN TECHNOLOGY ComputerLetter

Published by Technologic Partners

Volume 13, Number 28

August 25, 1997

Hide and Seek

New technologies that give an expanded role to client-side software may make the task of finding what we need on the Web a lot easier

LIBRARY

AUG 25 1997

LATHAM & WATKINS
SILICON VALLEY, CA

A friend who learned his navigating skills in the South Pacific during the Second World War once gave us his opinion of Columbus's skills as a navigator: Columbus wasn't quite sure where he was going, didn't know where he was when he got there, and when he returned, didn't know where he had been. To our mind, this pretty well sums up the feeling we often get while navigating the World Wide Web in search of enlightenment or at least information. With the Web's content having expanded to an estimated two terabytes of data, the sheer magnitude and disorganization of that information leads us to worry that no amount of searching by conventional means will uncover everything we might find useful.

We know that we aren't alone in this frustration. A recent survey done at Georgia Tech found that nearly a third of Web users are unhappy with their ability to find information on the Web. Search engines return lists of hundreds, even thousands, of sites to any query, the majority of them irrelevant to our needs or of questionable authority. There is the additional problem, as we've pointed out previously, of the so-called "invisible" Web — pages that do not exist as separate files but are generated on the fly when a user makes a specific query. The *Amazon.com* and Barnes & Noble book sites, for example, are almost entirely built from databases that create pages at the request of individual users. The spiders from the major search engines never find these pages, which is why typing

(Continued on Page Two)

Private Profiles: Onyx Software

Testing the Microsoft style on a fast-growing market opportunity

It's become a popular cliché when the subject of *Microsoft* spinoffs comes up to refer to the growing population of Redmond-area startups as "the Baby Bills." It would seem to be a perfect marriage for the hard-sell '90s: *Microsoft*'s bottom-line mentality and business focus joined with youthful entrepreneurship. Certainly all those programmers and managers aspiring to greatness with financing fueled largely by Bill Gates's stock-option largesse would like to believe some of his drive and acumen has rubbed off. Indeed, we've seen some promising young companies sprouting around Puget Sound. Of those,

we'd nominate *Onyx Software* as the most eager to achieve something like *Microsoft*'s spectacular success by following its lead at every turn. The question is whether *Onyx* will carry its devoted support for the mothership further than is healthy for a well-rounded offspring.

(Continued on Page Six)

This Week

In Our Opinion

After the search, the rescue 4

NEW WAYS TO SEARCH

Continued from Page One

"Raymond Chandler" into a search engine will not list either of these sites in the top ten.

Another problem is that, by-and-large, searches are text-based. This ignores all of the multimedia data — video, still images, animation, and sound — that exist on the Web. If you only want to see the trailer for "Conspiracy Theory," you shouldn't have to download a pile of text reviews until you find one that has the video attached. The problem of finding digital media will be especially acute on corporate intranets as they become more widely used as repositories for presentations and training material using graphics, sound, and video.

Today's search engines don't do well trying to find multimedia files or text that's buried in a database.

They just want to help

We are heartened, however, by the number of efforts that have sprung up over the last year or so to wrestle with some of these issues. Not surprisingly, they come mostly from new companies that sense an opportunity to create a substantial business out of providing smarter ways to search in the same way that **Netscape Communications** and **Yahoo!** prospered by offering basic navigational tools. We aren't sure that they are likely to find an application category as fundamental as the browser or the Web search engine proved to be. But we're fairly sure that they are working on problems of considerable importance to a lot of people — usually a good starting point for an entrepreneur.

We'd include on that list — but not in this letter — the companies developing push technologies of one kind or another. While the efforts of **BackWeb Technologies**, **Diffusion**, **PointCast**, and others are helping to get us information we already know we want, they don't do much for the problem of getting to sites we aren't familiar with to collect data that may or may not be of interest. Similarly, we aren't dealing here with collaborative filtering and other so-called agent-based technologies, the

bailiwick of companies like **Autonomy**, **Firefly Network**, and **Net Perceptions**, which build profiles over time and then automatically deliver information they think is relevant to our needs. (For our discussions of those subjects, see "Push Comes to Shove," November 4, 1996, and "In From the Cold," September 23, 1996.)

ComputerLetter

Editor: Richard A. Shaffer
120 Wooster Street
New York, NY 10012
Tel: 212 343-1900 • Fax: 212 343-0915
e-mail: shaffer@technologicp.com

Managing Editor: John W. Wilson
P.O. Box 969
Larkspur, CA 94977-0869
Tel: 415 924-1274 • Fax: 415 924-0945
e-mail: wilson@technologicp.com

Senior Editor: Tiernan Ray

Senior Writer: Craig Canine

Contributors: Kathleen K. Wiegner,
Ken Andersen, David Brand

Art Director: Natalie Tilton

Copy Editor: Jennifer DeFeo

Fulfillment Coordinator: Mike McNamara

© 1997 by Technologic Partners. All rights reserved. **ComputerLetter** is published 40 times a year by Technologic Partners. Nothing that appears in **ComputerLetter** may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without the prior written consent of the publisher. The title **ComputerLetter** is a trademark of Technologic Partners.

Subscriptions are \$595 per year in the United States. Outside the United States subscriptions are \$695. For subscription inquiries, or to order back issues and reprints, contact Technologic Partners in New York.

The information and statistical data contained herein have been obtained from sources we believe to be reliable, but are not warranted by us. We do not undertake to advise you as to any change in the data or our views. Technologic Partners and its affiliates, partners, and employees, or members of their families, may perform services for, engage in business with, or own the securities of one or more of the companies referred to in **ComputerLetter**, or their competitors, and may engage in transactions with those securities.



Technologic Partners

Neither push nor agent technology deals with a fundamental fact of human nature: Goals and interests often change from day to day. Push technology is fine for those constants people may always want — entertainment, news, weather, and sports at home, or financial updates in the office. But it is less helpful if you need to know in the next half-hour what the top selling drugs were last year because you lack some crucial facts for a report. Agents have the same drawbacks. Ideally, what a lot of people want is to type in a query — without a vast knowledge of Boolean search techniques — and quickly end up at the best source of information, minus the false starts and dead ends.

Brewster Kahle's latest brainchild assists searchers by keeping track of previous visitors.

Is free cheap enough?

That's the goal being pursued by several companies that have surfaced recently. In some cases, they hope to achieve it by putting special search capabilities on the client side rather than on the server, as is typical with conventional search engines. All have a revenue model reminiscent of that pursued so effectively by Netscape and other Web pioneers: Give away the software to users and make money by selling advertising on the server or by charging Web sites for the right to use the technology.

But the companies are using their improved search techniques for very different purposes. **Alexa Internet** tracks the movements of earlier visitors to a site, providing a kind of guide to trails blazed in the wilderness; **Citizen 1 Software** and **Netbot** offer access to otherwise-invisible data on commerce sites; **NetMetrics** provides short previews of Web pages to narrow the search process; **Netword** is marketing a search engine with a free client component that searches on keywords registered by organizations and individuals and stored in a central server. An exception to the client/server approach is **Magnifi**, which addresses the problem of searching image, sound, and video files by giving Web servers the power to summarize and index them.

Breaking the trail

Of these newcomers, we think Alexa may hold the most promise for the broad search market. The brainchild of information

scientist Brewster Kahle, father of the Wide Area Information Server (WAIS), a forerunner of today's search engines, Alexa starts with the assumption that the best way to find the most relevant sources of information is to learn from the experience of others. Alexa's client software provides users with a toolbar that accompanies them from site to site. Based on information that has been gathered about the decisions made by previous visitors to a site, the toolbar offers suggestions about where to go next.

Say you pick a site to learn about kayaking. Alexa tells you how well other kayakers before you who were using Alexa liked the site (did they keep on clicking or exit immediately?) and what path they followed through the site. It also tells you who publishes the site so you can decide on the authenticity of the information. Sites are rated according to how often they are updated, the speed of the server, and popularity with users. To find other useful sites relating to kayaking does not involve returning to the search engine or following embedded hot links. Instead, Alexa continues to offer up a list of sites that other users have frequented, and users are encouraged to vote on sites and to recommend links. Alexa's pitch to advertisers is that advertising within the toolbar, which is always displayed on the computer screen, provides longer exposure than ads on individual sites.

Coming attractions

In another approach to helping the search process, **NetMetrics** provides a downloadable piece of software that works within the browser to provide short previews of each Web page as it is downloaded. The product automatically pulls out headlines or key text to provide a hierarchical outline of all the pages in the site and a preview of each section. In theory, this enables the user to decide quickly whether to move on to a more useful site or to delve deeper. **WebTurbo** can query six of the most popular search engines at once and remove any duplication. Search results can be organized along with the previews and saved in topic folders. We found the previews presented by **WebTurbo** were sometimes garbled and of uneven quality but were, in general,

In Our Opinion

After the search, the rescue

If there's any lesson all the new startups can draw from the bloody jockeying for pole position engaged in by **Excite**, **Yahoo!**, and the rest of the search crew in the past year and a half, it's that the business of finding things on the Web is not just about search. Nor is it about the technology used to index, sort, classify, and find things online. Instead, all of that technology was merely a beachhead from which to launch a variety of content programming sites, editorial products, and one-off Web services that add value to what quickly became a commodity.

Inevitably, the intriguing client/server technology, agent intelligence, and database savvy of **Alexa Internet**, **NetBot**, and the others will ultimately provide no more than a foot in the door, allowing these companies to develop more sophisticated content and service businesses. The merger of client-side software and back-end databases will enable them to refine

those services gracefully over time. That's important, because in the not-too-distant future, the basic sorting and searching capabilities of these "vertically oriented" services will be devalued as the Internet evolves into a more sophisticated distributed network of directory servers, mobile agents, and robust metadata repositories.

As the software infrastructure of the Net gets smarter, online power plays such as **Netword's** gambit to own the very words we use, and thereby intrude into our daily browsing, will fade away. Such schemes cannot possibly scale to encompass an Internet growing by leaps and bounds. Attempts to control the way we surf will yield to the need to assess — and reassess — the fundamentals of selling software, packaging content, and luring advertisers. That's a tough race to run, but at least these companies are at the starting line. □

helpful in weeding out the obvious clinicians. We're told that the **NetMetrics** — which launched **WebTurbo** in July — expects to have 25,000 users by next month, at which time it will begin actively soliciting advertisers.

Taking a cue from **America Online's** "Keyword" searches, **Network** enables site owners to create their own "Networks" for specific content locations such as an annual report or company background. Users can then just type in these words rather than a lengthy URL to reach their destination by way of the **Network** server. **Network** recently signed up **Progressive Networks**, that pays a monthly fee which depends on how many **Networks** it registers. Individuals who want their own **Networks** for their home pages pay \$1 a month.

All wrapped up

Netbot, using research into intelligent agents at the University of Washington, has come up with a client/server search assistant that we wouldn't really classify as an

agent. **Netbot's** software called **Jango**, a shopping application that is the first in a series of application-specific helpers, allows the user to search multiple sites simultaneously for the product information stored in databases that other search engines can't penetrate. **Netbot** does this by endowing the client with natural-language technology that can take a query and figure out the exact terminology for what is being searched. Then the **Netbot** server routes the query to the appropriate sites, talking to them through an information adapter, also known as a wrapper, that must be written for each site to be searched. Future wrappers might be developed for job search, news retrieval, or white pages applications.

At last, some service

Jango's task is to help online shoppers locate and purchase goods from online merchants by searching the appropriate sites and then returning with product reviews, specifications, pricing, and other

Netword hopes you'll go through its Web site to get to everything on the Internet.

shopping details. Users can buy only from a few categories of goods at present, including books, wine, flowers, and computer hardware and software, because Netbot is focusing on only the top five or ten retailers in each category as well as information-based sites such as newsgroups for reviews. The company doesn't charge either users or commerce sites for the privilege of using Jango, hoping to live on advertising revenues.

For newcomers to online shopping, Jango seems useful in helping to find the places where books or flowers are sold and providing some comparison pricing. But veteran shoppers probably already have their favorite online stores and don't spend a lot of time comparative shopping. We also note several complaints, in that section of Jango's site devoted to user feedback, about the software missing obvious sites and the limited number of sites Jango brought up. Still, Netbot's notion of creating a more powerful vertical search capability is certainly one way to chop the Web into manageable chunks.

Getting specific

Like Netbot, Citizen 1 Software offers access to invisible Internet databases through a piece of free downloadable software that sets off parallel searches. But rather than just get a report, such as Netbot provides, users can view one

retrieved document while others from different providers load in the background. Along with its general-purpose version, Citizen 1 also markets industry-specific versions, initially for professionals in the pharmaceutical, device and diagnostic industries, called Citizen 1 HIT. Searching on diabetes, for example, will turn up information about everything from clinical trial results to ongoing research — all from within a single application. Citizen 1 makes searching more convenient by organizing databases into categories. Users choose which to search based on brief descriptions, and the chosen databases are searched simultaneously. Such refined capability doesn't come cheap, however. Citizen 1 HIT is priced at \$1,500 per seat.

Picture that

Even without a client component, the new technology can help Web servers at big sites do a better job of responding to queries. Netbot, for example, is preparing a server version of Jango. And that is the primary strategy for Magnifi, which was cofounded by Eric Hoffert, one of the original developers of Apple Computer's QuickTime multimedia technology. Rather than forcing users to do a text search and download a page to find out whether there is any sound or video, the Magnifi server installed on a site provides

Jango's appeal may be limited by its scant selection of stores.

Who's Where

Private companies developing alternative Internet search technologies

Company	Phone	Headquarters	URL	Products
Alexa Internet	415-561-6900	San Francisco, CA	www.olexa.com	Navigational aids based on usage patterns
Citizen 1 Software	415-882-9404	San Francisco, CA	www.citizen1.com	Industry-specific search software
Magnifi	408-863-7200	Cupertino, CA	www.magnifi.com	Server that allows indexing of multimedia for searches
Netbot	206-522-7800	Seattle, WA	www.jongo.com	Web shopping assistant
NetMetrics	415-248-2333	San Francisco, CA	www.webturbo.com	Software for browsing at the preview level
Netword	703-528-9600	Arlington, VA	www.netword.com	Keyword system for Internet searches

Copyright © 1997 by Technologic Partners

a preview of an image, video, or audio clip. From the perspective of a site owner with a lot of multimedia in the database, Magnifi provides a way to organize this content so that it is easily and quickly searchable and therefore likely to pull more traffic. CNN Interactive, Hollywood Online, and PBS Online have installed Magnifi's server and report increased traffic as people discover that they can do better searches. While this should lead to higher advertising rates, Magnifi plans to make its money selling server software at \$20,000 a pop rather than to seek a share of the ad billings.

Call to arms

The big search sites can't afford to sit still when it comes to solving some of the problems addressed by these startups. The client-side companies, in particular, pose to be a threat because they can take users directly to a site without having to check in with a search site. Then there is

the increased competition from well-positioned newcomers taking direct aim at their core business. CNET, for one, plans to debut its Snap Online service before year-end, featuring a new search engine that will not only let users search globally but enable them access to the databases of some 75,000 sites selected by a Snap team and updated every twelve hours. Microsoft is also expected to introduce a search service, currently code-named Yukon.

Considering the survival-of-the-fittest urges driving all the large search companies, we think it will be difficult indeed for any of the newcomers to hold more than a temporary technological lead. But the need for quick and efficient access to the Web's vast resources is so compelling that at least some of them will, one way or another, become an integral part of the search process. □

A Microsoft-centric strategy for customer management software

(Continued from Page One)

In terms of revenue growth, Onyx is already well ahead of its role model. It took Microsoft five years to get to \$8 million in revenue after Paul Allen and Mr. Gates launched the company in 1975. Onyx, founded in 1994 by three Microsoft alumni, was profitable from the first year and had sales last year of \$9.7 million, a gain of 350% over 1995. This year President and CEO Brent Frei and Executive Vice President Brian Janssen are confidently predicting a further doubling of revenues. With customers now numbering 140 — 60% of them in the software industry — Onyx clearly has staked out a prominent position in the highly competitive market for customer management and information software.

Big Green attack

That niche was a natural for Mr. Frei and Mr. Janssen, Dartmouth football teammates in the late 1980s, who had developed a customer management database for Microsoft's international subsidiaries. When Mr. Janssen was asked by Spry,

CompuServe's Internet software division, to improve its databases for marketing, sales, and customer service, he turned first to Mr. Frei and then to Todd Stevens, another young Microsoft colleague. They started Onyx in the basement of a suburban home east of Seattle with a \$150,000 personal investment and set out to develop the best possible customer support product based on Windows NT and, later, on Microsoft SQL Server.

Onyx entered a field that had been pioneered with Unix-based client/server products by companies such as Aurum Software, Clarify, Scopus Technology, Siebel Systems, and Vantive. The market was already growing briskly; Gartner Group's senior analyst Wendy Close says the overall use of technology for selling-related functions is rising by more than 50% annually, although the category is difficult to define since it involves field sales and service, telemarketing, external help desks, and support for third-party resellers. Analysts call it technology-enabled selling, customer interaction

Onyx is already ahead of Microsoft's early growth pace.

management, support services, and various other labels. Ms. Close estimates that salesforce automation alone generated \$2.7 billion in sales worldwide last year. Counting customer service and support, the total for these overlapping product lines reached a whopping \$3.4 billion.

customers pay about \$2,500 per user, compared to as much as \$5,000 for competing applications, which typically are sold on the basis of separate pricing for individual modules. On the other hand, Onyx's all-in-one approach can be more expensive if the customer needs only a couple of functions.

What drives the growth of this market is the high cost of winning and keeping new customers.

Finders and keepers

The engine driving this growth is the high cost of winning and retaining new customers in the super-competitive sales environment of the 1990s. The *Harvard Business Review* has estimated that U.S. companies are now suffering an annual customer turnover of 50%. The big challenge for management is to reverse this tide of customer defections and at the same time build new customer relationships by increasing sales and support effectiveness.

Onyx's Mr. Janssen points out that it is ten times more expensive to acquire a new customer than to sell to an existing client. Thus, he argues, it's important that every employee dealing with customers be able to get information quickly — something that might be hard to do when customer data is scattered across many applications and databases. One of Onyx's first customers was *Visio*, a producer of business graphics software, which consolidated several applications into the Onyx Customer Center package to provide a single point of contact for 55 Visio employees authorized to use and update customer information. For this convenience, Onyx's

Follow the leader

While the product niche seems promising, it may be Onyx's relationship with Microsoft that is most vital to the company's success. For one thing, Mr. Frei and Mr. Janssen have followed the iron-clad alumni rule: Don't compete with the alma mater, but do develop a market that can ride on Microsoft technology. Onyx went into customer management software in part because Microsoft doesn't make it, and the company has developed exclusively for Microsoft BackOffice. Onyx software runs only on Windows NT and Microsoft SQL Server, a self-imposed limitation that has the potential advantage of reducing time to market for upgrades and new products by eliminating the need to create multiple versions.

Mr. Janssen likes to brag that the company has avoided an R&D expense "matrix of pain" by not supporting multiple systems, and we agree that a single-platform focus should lead to lower expenses. But we think there are serious risks as well. Some analysts view the exclusive commitment to Windows NT as

ComputerLetter

BUSINESS ISSUES IN TECHNOLOGY

Published by Technologic Partners

Name _____
 Title _____
 Company _____
 Address _____
 City _____ State _____ Zip _____
 Phone _____ Fax _____
 E-mail _____

- ☐ **Sign me up!** Enclosed is my payment of \$595 for a one-year subscription (40 issues) to *ComputerLetter*. Outside the U.S., please add \$100 for postage and handling.
- ☐ **My satisfaction is guaranteed!** If, for any reason, I am not satisfied with the newsletter, I will receive a full refund on all unmailed issues.

- ☐ Check enclosed (payable to Technologic Partners)
☐ American Express ☐ Visa/Mastercard

Signature _____

Card no. _____

Exp. date _____

Return to:

Technologic Partners • 120 Wooster Street • New York, NY 10012
 212-343-1900 • Fax: 212-343-1915 • computerletter@technologicip.com

At a Glance: Onyx Software

Address: 330 120th Avenue NE
Bellevue, WA 98005
Telephone: 425-451-8060
URL: www.onyx.com
President & CEO: Brent Frei
Business: Customer management software
Financing: \$11 million
Investors: Foundation Capital,
Technology Crossover Ventures
Employees: 175
Revenue: \$9.7 million in 1996

As the boom in customer service applications plays out, the industry is likely to see some consolidation.

too limiting, in part because the scalability of NT is not yet proven for large organizations. If Onyx wants to graduate from selling primarily to medium-sized companies, says Gartner's Ms. Close, it will be forced to support Oracle database servers and Unix hardware.

Comes the Revolution

Other analysts also fault Onyx for producing software that has only limited tools to customize and configure the application for individual needs and for continuing to operate with a 16-bit architecture when much of the world has shifted to 32-bit technology. The company is also faulted for being too focused on the needs of customer service and support departments and for not being sufficiently attuned to field sales organizations that require software for setting sales goals and forecasting. Onyx says it does have such software, called Revolution, under development and promises that when it is shipped early next year the product will be fully ported to a 32-bit environment.

Where Onyx seems most like its giant neighbor is in an aggressive and innovative approach to sales and marketing. As one example, the company junked the usual mix of direct mail, advertising, and trade shows when it found that its World Wide Web presence was the most efficient and successful marketing channel. Now the Web provides over 30% of sales leads and has helped give Onyx a reputation for responsiveness.

In the meantime, Onyx's competitors are tearing ahead. Although this is an industry without a standard-bearer, there are leaders in specific sectors. Siebel, for example, is emerging in the vanguard of sales opportunity management systems, Dun & Bradstreet's Sales Technologies subsidiary in applications for a healthcare products sales force, and Austin-based Trilogy Development Group in product configuration systems.

Our venture friends believe that the current market boom will continue for about the next five years, with consolidation slowly taking place as the front-office sales and service applications are integrated with back-office manufacturing and financial software. This process has already begun with the recent acquisition of Aurrum by Baan, the Netherlands- and California-based business software company. We imagine that Oracle, PeopleSoft, and SAP might be thinking along the same lines.

Please be more specific

In such a market, newcomers like Onyx will have to find a way to stand out. Despite the many claims made by all the entrants in this developing market space, there remain doubts that the vendors are offering much more than a general solution to the highly individual problems companies face in maintaining and strengthening relationships with their customers. Paul Koontz, a partner at Foundation Capital, an Onyx investor, notes that the market is so big that the software companies have not yet been forced to develop any true vertical alignment or to give their products the flexibility needed to adapt to differing needs.

What clients really want, say analysts, and what will take much more development, is a product that is rich in functions and with tools that are readily adaptable, so that no sales force is locked into just one way of doing things. Onyx might well be among the first to achieve this. The company, we are told, is working on ways to make its product less monolithic and more flexible, a direction it must ultimately take if it is to attract larger organizations. Onyx knows, like the godfather it emulates, that it can't afford to have any ideas cast in stone. □

Why should your Web site be *DIGEX* driven?

internet
news.com

product news

Alexa Highlights Web Site "Trustmarks"

[August 25] Web navigation device Alexa Internet will now display a privacy "seal of approval" as designated by TRUSTe for those Web sites promoting secure online privacy.

TRUSTe is a nonprofit organization that assigns "Trustmarks," or logos, that when clicked on provide users with information on what kind of information a Web site collects, what it does with that information, and with whom that information is shared.

Alexa Internet is a search tool that includes detailed information on Web sites being visited, including who owns the site, where the company is physically located, how many pages it contains, and how "fresh" the site is, or how often a site updates its content.

Launched in mid-July, Alexa Internet is a toolbar that resides below a user's Web browser window. Users will be able to find out whether a Web site displays a Trustmark upon opening the "Where Are You Now" window in the toolbar.

"With TRUSTe, our users now receive additional information regarding the way a site handles personal data. For example, an individual interested in purchasing clothes on the Internet could quickly view which sites adhere to TRUSTe's privacy standards before they spend the time to go into the site and judge whether they should do business with this site," said Brewster Kahle, president and co-founder of Alexa Internet.

The Alexa toolbar is compatible with Windows 95 and Windows NT operating systems, and can be downloaded for free from Alexa's Web site.

<http://www.internet.com>

by [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

PCWEEK

AUGUST 18, 1997 • Volume 14, Number 33

BREAKING STORIES www.pcweek.com

Alexa's Theory of Relativity

REVIEW: Filtering, analytical algorithms link to Web sites—relevant or not

BY JIM RIZPOLA, PC WEEK LABS

A PRESENTING DATA PROGRAM FROM A SAN Francisco startup helps Web users find useful information by applying collaborative filtering technology to the entire Web.

Alexa, from Alexa Internet Inc., provides a series of links in Web sites that are related to the one the user is viewing. PC Week Labs found it very straightforward to use, letting us view basic information such as a Web site's owner, its popularity (based on number of visitors) and the freshness of its information (based on how often it is updated). We also could vote on whether or not a suggested site was useful. In some cases, the Alexa beta led us to useful related sites that we didn't know

other vendors to help commercial sites create personalized interfaces with visitors. Alexa is among the first to apply this technology to the entire Web. Currently, because the beta is new and limited in scope, Alexa creates the list of related links for a site by monitoring where users went after viewing that site.

The beta release let us vote on a suggested site's relevance, but it wasn't clear if Alexa included our vote in its determination of related sites. In the future, the program will determine links based partly on users' votes, company officials said.

Creating links based on users' Web crawling yielded hit-or-miss results in tests. When we viewed the PC Week site, related links included some worthwhile sites,

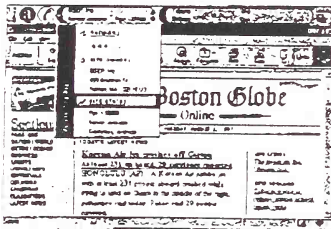
but also some sites whose relevance was questionable, to say the least (the National Computer Television Online site also was listed, for example). Other tests provided links with more relevance to the test site, however.

Speed and freshness proved to be highly subjective and not very useful in tests. For example, server speed is currently based on how long it takes to retrieve pages from a site to Alexa's offices in San Francisco. Although this might be useful to Alexa employees, it means almost nothing to users, especially those in other parts of the country or in other countries.

This measure of speed and freshness also caused Alexa's own site to appear much more powerful than such huge sites as Netscape Communications Corp.'s or Yahoo Inc.'s locales—Alexa scored a Fast rating compared with these sites' Average ratings in our tests.

A site's freshness is determined by the average age of pages on the site, with fresh sites having an average age of less than six months; stale sites an average of more than two years, and all the rest rated as average in age.

This led to very misleading scores in our tests—sites that are updated several times per day got average scores because they have very good archives, whereas new sites that hadn't been updated in a few months scored as fresh. We think it would be more useful to simply rate sites as fresh (less than 6 months old) or stale.



Alexa provides information about Web sites and links to related sites.

or, in other cases, however, sites that Alexa considered related were only vaguely related or were irrelevant to us. The basic site information that Alexa provides also needs some work.

The program, which is set to ship next month, will be free. The company hopes to finance the program by selling advertisements that will appear in the tool bar. Alexa works with Microsoft Corp. and Netscape Communications Corp. Web browsers running under Windows 95 or Windows NT.

Crawling through fields

Alexa data is gleaned from a combination of collaborative filtering and analytical algorithms applied to frequent Web crawling.

Collaborative filtering technology is typically used on specialized electronic commerce sites such as Amazon.com and is based on the premise that users with similar interests will like the same things. For example, if a user liked a certain fiction book, King might be the way to suggest a related book. The way site might recommend a Clive Barker book that was liked by other users who enjoyed the King book.

Collaborative filtering technology is used by products from Firefly Network Inc. and

other vendors to help commercial sites create personalized interfaces with visitors. Alexa is among the first to apply this technology to the entire Web. Currently, because the beta is new and limited in scope, Alexa creates the list of related links for a site by monitoring where users went after viewing that site.

When we viewed the PC Week site, related links included some worthwhile sites, but also some sites whose relevance was questionable, to say the least (the National Computer Television Online site also was listed, for example). Other tests provided links with more relevance to the test site, however.

Speed and freshness proved to be highly subjective and not very useful in tests. For example, server speed is currently based on how long it takes to retrieve pages from a site to Alexa's offices in San Francisco. Although this might be useful to Alexa employees, it means almost nothing to users, especially those in other parts of the country or in other countries.

This measure of speed and freshness also caused Alexa's own site to appear much more powerful than such huge sites as Netscape Communications Corp.'s or Yahoo Inc.'s locales—Alexa scored a Fast rating compared with these sites' Average ratings in our tests.

A site's freshness is determined by the average age of pages on the site, with fresh sites having an average age of less than six months; stale sites an average of more than two years, and all the rest rated as average in age.

This led to very misleading scores in our tests—sites that are updated several times per day got average scores because they have very good archives, whereas new sites that hadn't been updated in a few months scored as fresh. We think it would be more useful to simply rate sites as fresh (less than 6 months old) or stale.

Alexa 1.0 Beta

As a free add to a user's Web browser, Alexa shows some promise for helping users find sites that other users found to be useful or interesting. However, the beta could use some improvements to its interface and to how it determines site information before it's ready for heavy use.

- Helps users find sites related to those they are browsing.
- Information provided on sites can be misleading; related links don't always make sense.

Send Internet fax to San Francisco (415) 333-0000
www.alexainternet.com
 or by e-mail: alexainternet@alexainternet.com

If Alexa can improve site information it would be useful if that information were supplied for recommended links, as well. This would, for example, let users know a site's rating and freshness before they actually accessed that site.

Alexa includes a chat interface that lets users communicate in real time with other Alexa users. We liked that a user had to know another user's address in order to chat, because this saved us from being bothered by unwanted chat requests.

Alexa's archive capability feature was not fully functional in the beta we tested, but this feature will eventually allow users to access an online archive, maintained by Alexa, of pages that are either temporarily not available or have been disabled. ■

Insight auction

computers, hardware and software

\$35

c|net

CNET.COM

home

[Free newsletter!](#)[NEWS.COM](#)[Radio](#)[MENU](#)[Reviews](#)[Hot list](#)[Just in](#)[All comparisons](#)[CD-ROM central](#)[Best of the Web](#)[Features](#)[How to](#)[Digital life](#)[Techno](#)[Events](#)[BUILDER.COM](#)[Browsers!](#)[Personalities](#)[CNET TV](#)[The shows](#)[The hosts](#)[Airtimes](#)[CNET studios](#)[Community](#)[About CNET](#)[Feedback](#)[Member services](#)[Posting](#)[Current poll](#)[Conferences](#)[CNET jobs](#)

/theweb/ 0023499961100

Download of the week

Alexa

This is one of the most interesting and useful Internet tools we've ever seen. The premise is simple: as you surf the Web, a toolbar appears on the bottom of your Web browser. This toolbar contains several buttons, each leading to information about the site you're visiting. You can check a site's quality rating (1 to 5), the speed of its connection, how many pages it contains, and the relative "freshness" of its material. You can also find out who owns the domain name for each site you visit, and exactly where the company is located—useful if you like to dash off letters to your favorite sites' owners, but can't ever seem to find their email addresses.

But the highlight of Alexa is its recommended links button. For every site you visit, Alexa will suggest up to ten other sites with similar subject matter. The recommendations are sometimes way off base, but they're far more intelligent than the usual robot-generated responses. For example, a visit to [CNN](#) gives you direct links to competing news organizations; a trip to [The Dilbert Zone](#) can send you down a pathway to other extremely popular newble sites like the [White House](#); and a visit to a catalog of porn sites will—rather surprisingly—give you lots of wicked suggestions.

You can grab a beta copy of Alexa today. You need a 486 or better PC running Windows 95 or Windows NT, and Netscape Navigator 2.0 or Internet Explorer 3.0 (or higher). Just click the link below to get started:

[Download Alexa](#)

Start to index

[Past downloads](#)[Screensavers](#)[Fighting, fish, and freaky psychedelica](#)[Deathmatch Maker](#)[Create your own Quake levels](#)[f.Search](#)[Conduct Net searches from within other apps.](#)[CleanSweep 95](#)[Flush out your Windows 95 PC.](#)[Net.Medic](#)[Why's your connection so slow?](#)[Paging](#)[Paging, right on your PC screen.](#)

recording technology was primitive. "When it comes to a point where users have a camcorder recording the Net, then the archive won't be worth it," he says.

In Building 116, the archive is stored in a digital-tape library that looks like a vending machine. The tapes currently have the capacity to hold 20 terabytes of data in all, about as much information as is in the Library of Congress. So much content is being added to the Internet that the archive grows by about a terabyte of data each month. The data-mining computers are able to adjust their site visits to concentrate on those that change most frequently. They will come upon a site, however, only if Alexa users have visited it, if anyone else on the Web has linked to it, or if it is listed with a directory service.

While the archive has been able to keep up with textual information on the Internet -- it is complete from October 1996 to the present -- the effort to collect images is running a few months behind.

Through Alexa, the archive receives about 14 requests for old pages every second. Not bad, its founder says, when one considers that only about 9,000 people visit the San Francisco Public Library on an average day. Alexa also helps researchers by listing facts about the site they're visiting: the address of the individual, company, or other organization that owns the server on which the site is located; how many people have visited the site; how frequently the site is updated; how fast its computers are; and how many pages the site contains. Alexa also allows users to vote for their favorite sites and keeps a running total on each site.

Mr. Kahle says Alexa does not keep individual statistics on its users. Although the search engine tracks the paths of users as they jump from site to site, it does not record users' names. "We don't care who you are," he says. "We just care what path you take."

Mr. Kahle dreams that Alexa could become as popular -- and as profitable -- as search engines like Yahoo! and Excite. And the Internet Archive, he says with enthusiasm, could become part of a large research library, although he's not sure how. "I don't think about the details," he says. "That's why we're doing something now that others thought was impossible, or even crazy."

Copyright (c) 1998 by The Chronicle of Higher Education
<http://chronicle.com>
Date: 03/06/98
Section: Information Technology
Page: A27

least some of the top sites," Mr. Kahle adds.

Still, one needs a traditional search engine or a specific Web address to get started, Mr. Kahle acknowledges. And Alexa, unlike the Alta Vista search engine and others, can suggest linking only to entire Web sites, not to specific pages within them.

"I don't think of it in the same way as a search engine -- it's a supplement," says Bruce Livett, a reader and deputy head of the biochemistry and molecular-biology department at the University of Melbourne, in Australia. "Alexa gives you relevant sites in the general sense, sites that you sometimes miss because other search engines depend on specific keywords you enter."

Dr. Livett, who has been using Alexa since October, surfs the Web to keep up with the research work of colleagues around the world. "It's competitive work, and I need to know what they're doing." Alexa, he says, has alerted him to research sites that did not turn up in searches using Excite and Anzwers, a search engine designed for Web users in Australia and New Zealand.

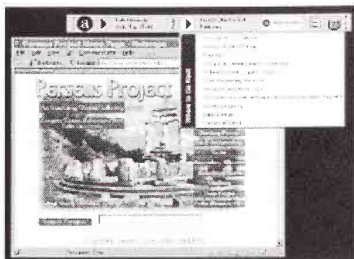
Part of Alexa's appeal, he says, is access to the Internet Archive. When Alexa users get a dreaded "404 -- file not found" error, they can click on a button on Alexa's tool bar and pull up the missing page from the archive. Using the archive, Mr. Livett found an audio interview he needed that had been removed from a Web site.

Alexa is "an immediate use for the archive," says Mr. Kahle, adding that the archive is the component that he expects will eventually separate his search engine from the rest of the pack. Alexa also offers a direct link to the *Encyclopaedia Britannica* Web site, allowing users to retrieve reference information without leaving the Web page they are viewing at the time.

As more people begin to use Alexa and the archive, however, tricky questions about copyright and privacy have begun to crop up. An e-mail discussion list for Web publishers recently included a heated debate about copyright issues surrounding old newspaper articles that are part of the archive.

The data-mining computers skip Web pages that require passwords, as well as Web sites protected by the Standard for Robot Exclusion, which blocks search engines from copying pages or directories. Still, some Web publishers said in the e-mail discussion that Alexa officials should be asking on-line newspapers and journals if they want to be part of the archive, instead of forcing them to block Alexa from copying pages.

How the archive will be used in the long term is not clear. Mr. Kahle often mentions the early days of television, when programs were broadcast live and



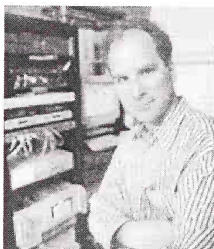
For example, from the "Perseus Project" (<http://www.perseus.tufts.edu>), a site with an extensive collection of ancient Greek texts in translation, Alexa points the user to sites about classicists and Mediterranean archaeologists at the University of Michigan, sites about publishers and journals available electronically, sites about Hellenistic linguistics, and to "Project Gutenberg," an Internet producer of free electronic texts.

Since October, more than 200,000 people have downloaded Alexa. The service, which Mr. Kahle hopes will soon be fully supported by advertising revenue, is not yet turning a profit. But by the end of the year, he expects it to have a million users. Eventually, he hopes to use the profits from Alexa to finance the gathering of data for the archive.

The advantage of Alexa as a search engine is that it "attempts to be an objective source" for people seeking information. Where conventional links are chosen by a page's creator according to what the creator knows and prefers, Alexa also brings other Web users' knowledge and preferences to bear. The sites recommended in a given search sometimes change, depending on the surfing patterns of Alexa users.

The system has its oddities. If users frequently traveled from the "Perseus Project" to, say, *The New York Times*, the newspaper could be added to the top-10 list of an Alexa user looking at the Perseus site, even though the only thing the two sites have in common is their users. In fact, such a situation has already occurred. From the Perseus site, Alexa suggests -- based on other users' habits -- visits to the sites of Franklin and Marshall College and Bates College. Alexa officials say students at the two colleges probably use the Perseus site in their classes.

"It's sometimes random and not always perfect," Mr. Kahle says. "But if researchers use a traditional search engine, they may miss some of the best sites." A search engine such as Excite, using the keywords "Greek texts" to find sites related to the "Perseus Project," turned up 268,057 matches. "With Alexa, you're bound to hit at



Fred Mertz for The Chronicle

Brewster Kahle started the non-profit Internet Archive after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million.

Just as the Internet has allowed all kinds of ordinary people to become their own publishers, it has allowed a computer scientist named Brewster Kahle (*left*) to create the non-profit Internet Archive.

Mr. Kahle decided to save the Internet's contents for posterity after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million. The system, which he invented, makes it easier to search electronic data bases.

Why an archive? "We need to preserve this heritage," says Mr. Kahle, an affable and enthusiastic 37-year-old who is a graduate of the Massachusetts Institute of Technology. "Or one day, digital anthropologists will wonder if we ever learned anything from the history of other inventions. Remember, nobody recorded television in the early days."

Once he started collecting all that information, Mr. Kahle says, he realized how difficult finding things on line was becoming, with the number of Web sites doubling every six months even as other material falls into neglect. So he set about creating a Web search engine using the technology he developed to manage the massive amounts of data he was collecting for his quirky history project.

The result is Alexa, a search engine operated by Alexa Internet, the for-profit company that is part of the Internet Archive. "This will change the way that researchers use the Internet," Mr. Kahle says.

Alexa is software that can be retrieved free from the company's Web site (<http://www.alexacom.com>) and added to a Web browser. Unlike other search engines, such as Yahoo! and Excite, it doesn't rely on word searches. Instead, it watches where its users go on the Internet, and then records that information in a central data base. Based on that information, Alexa can tell a user the most popular paths that other Alexa users have taken from the site the user is visiting at a given time.

It also can suggest other sites offering related material. The top 10 sites pop up in a thin, gray bar near the browser (*see below*) and change as the user moves from page to page.

INFORMATION TECHNOLOGY

March 6, 1998

In Attempting to Archive the Entire Internet, a Scientist Develops a New Way to Search It**Non-profit group uses 'data mining' in effort to preserve World-Wide Web content**

By JEFFREY SELINGO

SAN FRANCISCO

For 110 years, Building 116 served unobtrusively as a general store and as quarters for non-commissioned officers stationed at the Presidio, the U.S. Army base south of the Golden Gate. Now that the base has been decommissioned, Building 116 stands out.

Its red-tile roof and cream-colored shiplap siding have been restored, for one thing. And it's one of the few buildings still in use on the 1,480-acre property, which became a national recreation area in 1995. What makes Building 116 unique, though, is what's inside -- a virtual preservation project that aims to create a complete archive of the Internet.

In a back room, a buzz comes from giant computers that are "data mining" the World-Wide Web and Usenet discussion groups, finding and recording pages of the computer network on a nearby digital-tape machine. The computers take complete snapshots of the Web every two months, allowing users to find pages long after their owners have taken them down and let their hyperlinks lapse. So far, the archive has compiled eight terabytes of data -- the equivalent of 800,000 books -- and has recorded at least three snapshots of more than 500,000 Web sites.

Inside those virtual books is the patchwork history of ordinary people: pages of college students long graduated; Web sites of political campaigns since forgotten; early, awkward versions of sites that are now well-known; infamous sites that held our attention for weeks, such as the Heaven's Gate cult's page.

ALSO SEE:

[The World-Wide Web site of the Internet Archive, and a new Internet search engine that can be used to find material in the Internet Archive](#)

[Search The Chronicle](#)

[The current Chronicle](#)

[Copyright & reprints](#)

[Feedback](#)

Why not?

PUSH!

CNET.COM

home

[Free newsletter!](#)
[NEWS.COM](#)
[Radio](#)
MENU
[Reviews](#)
[Hot list](#)
[Just in](#)
[All comparisons](#)
[CD-ROM central](#)
[Best of the Web](#)
Features
[How to](#)
[Digital life](#)
[Techno](#)
[Events](#)
BUILDER.COM
Browsers!
Personalities
CNET TV
[The shows](#)
[The hosts](#)
[Airtimes](#)
[CNET studios](#)
Community
[About CNET](#)
[Feedback](#)
[Member services](#)
[Posting](#)
[Current poll](#)
[Conferences](#)
[CNET jobs](#)
Alexa

.

Alexa shows the way

 by Christopher Lindquist
(8/13/97)

just in
[back to index](#)

Search engines, agents, push clients--everyone wants to tell you where to go on the Web. Unfortunately, most solutions spit out a dry list of links based on keywords. Alexa Internet's

Alexa service takes a more democratic approach. And it's more than just a link list; it's a full-blown information service that gives you the skinny on nearly every site you visit.

The Alexa toolbar floats on your screen. As you surf, the program retrieves a variety of site information from Alexa Internet's servers. Go to any site, and Alexa reports the name of the company that owns the domain name, the server's relative speed, and how many pages exist on the site.

You also get a list of alternative site links. Rather than simply querying a search engine, Alexa invokes a variety of methods to guide users. Alexa's servers constantly crawl the Web and search page text to find similar content. The servers also analyze patterns of where people travel on the Net. Add a site to your link list, and Alexa notes the new site and gives it a high rating. If enough people add the site, it may eventually get onto the toolbar's default site list. The folks at Alexa Internet can also manually insert sites into a list to bring interesting new locations to the front. Finally, Alexa sells advertising space on its lists in the form of tiny color banners that also act as links.

Alexa's unique search approach often produces interesting results. For instance, a quick AltaVista search on *woodworking* called up the usual unfathomable 62,000 listings. However, after we made a quick visit to a favorite woodworking hobbyist site, Alexa produced links to four similar--and useful--pages we had never seen before.



Stats summary.

Resources[Software central](#)[Glossary](#)[Help](#)**CNET SERVICES**[NEWS.COM](#)[BROWSERS.COM](#)[BUILDER.COM](#)[GAMECENTER.COM](#)[SEARCH.COM](#)[DOWNLOAD.COM](#)[SHAREWARE.COM](#)[ACTIVEX.COM](#)[MEDIADOME.COM](#)**MARKETPLACE**[BUYDIRECT.COM](#)[Specials](#)[CNET store](#)[How to advertise](#)

That said, things don't always work as expected. Pages focused on words with multiple meanings can produce odd results (for instance, a church mission link attached to a site touting mission-style furniture). Some smaller, less-frequented sites don't produce any links at all. And Alexa's algorithms for determining a site's average freshness of content need some tweaking (they're biased against large sites).

Once you learn to navigate the sometimes cryptic toolbar interface, you'll be able to contact Alexa's tech support, enter a real-time chat with other Alexa users (though chat features were rather limited in our beta software), check a site's popularity, rate sites with a thumbs up or down, and dig into financial info about a site's parent company. If the page you want no longer exists on its original server, you may be able to retrieve the page from Alexa's extensive archives.

If you've been looking for a tool that helps provide some context for the sometimes chaotic Web, give Alexa a try.

(A limited beta is available now. The final version is due at the end of this summer.)

**facts**

Alexa

Direct price: free

Alexa Internet, 415/561-6900

PC 486, 3MB disk space, Windows 95 or Windows NT, Navigator 2.0 or Internet Explorer 3.0

[Back to top](#)

Copyright © 1995-97 CNET, Inc. All rights reserved.



NET INSIDER

[Insider Profile](#)

[Profiles](#)

[Insider Scorecard](#)

[Discussion Threads](#)

[Audio](#)

News

[The Insider's Take](#)

[The Scoop](#)

CMPnet RESOURCES

[Home](#)

[Site Map](#)

[Search](#)

[Ad Info](#)



Does your business

Free Trial

Microsoft
Site Server,
Enterprise
Edition

[Insider Profile Top](#) [Profile](#) [Interview](#) [Audio](#)

INSIDER PROFILE:

Brewster Kahle

[»read interview](#)

[»hear interview](#)

Job

Internet Archivist,
Entrepreneur



Home

Lives in one of 30
officers' homes, now
part of Presidio
National Park, in San
Francisco.

Education

B.S. in Computer
Science, Massachusetts Institute of Technology,
1982

Most recent accomplishment

Archived the Web on 2 terabytes of storage,
devised system to navigate Web following the
trails of where others have gone.

Track Record

Built Connection Machines at Thinking Machines
until 1992. Founded WAIS Inc. in 1992. Sold
WAIS to America OnLine in 1995 for \$15 million.
In 1996 left WAIS to form [Alexa Internet](#) and the
[Internet Archive](#).

Message of Optimism

"I think the technology we are building is really
trying to interconnect people better. If people
communicate better and faster, you can make the
whole organism smarter. That's what we're
shooting for."

Message of Pessimism

"It takes longer than you can possibly imagine."

Personal

Married, two sons.

Cache This

Brewster Kahle
CEO
Alexa Internet

- www.turtles.org/happen.htm#logger

This site tracks a turtle's migration from California to Japan in a way that opens science to students—an example of sensors getting integrated with computer networks in interesting ways. If this kind of thing catches on, then maybe life doesn't have to be a spectator sport.

- www.epic.org

The Electronic Privacy Information Center's aggressive but sane presentation of one of the trickiest and most important social issues happening in the technology sphere.

- pharmdec.wustl.edu/iglu/guiz/surrealism.html

The power of Web publishing—inexpensive, highly interconnected and equalizing—allows for sites like this one on surrealism.

View All Bookmarks Columns

b o o k m a r k s

[Home](#) | [Daily News](#) | [In the Know](#) | [People](#) | [Companies](#) | [Financial](#) | [Books](#)

Feel free to send E-mail to: [Ad Sales](#), [Editorial Feedback](#), [Subscriptions/Back Issues](#), or [Rugs](#).

[Subscribe to UPSIDE today!](#)

Copyright ©1997 Upside Media Inc. All rights reserved.
Powered by [Thunderstone](#).

WIRED

WI

CULTURE



WEB PAGE QUOTE



WEB PAGE QUOTE



✕ Net Surf: Alexa's New Navigation Service

by **Joey Anuff**

6:06pm 1.Aug.97 PDT "Net Surfing," as a concept, is an anachronism - a throwback to the days when the Web was all about motion, and rarely about rest. In the days when every page, personal or corporate, was a "homepage," the common denominator was the hotlist, and every Web site felt compelled to carve its identity via a list of links. It made sense - before server-pushes, Java, and stylesheets, the most sophisticated Web technology was the href.

It still is. Too much so, in fact. By now, almost everybody realizes that traffic is the true global currency. Outside of bona fide homepages and Yahoo, most off-site links are ads, paid for by people who need more traffic in order to sell more traffic. The surreal mathematics of trafficking in traffic may become better understood and standardized, but they won't go away. Which creates a huge opportunity, a chance to reintroduce the connectedness the Web was originally created to enable. Thus [Alexa](#) - a new navigation service conceived by Brewster Kahle, inventor of WAIS and co-founder of Thinking Machines.

Ironically, while reminiscent of quite a few collaborative agent filtering sites, search engines, and bookmark-management schemes, Alexa is neither a Web site nor a plug-in, but an application that works in cooperation with your browser. It functions as a toolbar that presents an active analysis of every site you visit - who owns it, how many sites link to it, how many pages are on it, and how popular it is amongst other Alexa users. More importantly, it provides multiple suggestions for where to go next - based on which sites are similar, which sites others have fled to, and which sites others have linked it to. And, of course, who has paid Alexa to be included.

But don't fault them for the slight intrusion of commercialism. That's a given. What's more alarming is the fact that the ploy works, more or less. Some might be impressed by its instant-messaging add-on or its archive service, which strives to eliminate 404s forever, but these are merely equivalent to leather interiors and a rear spoiler, entirely tangential to the quality of the ride. Even in its infancy, with its current surf suggestions based not on user feedback but on more mundane technical comparisons, Alexa's pointers often make no sense at all, but are precise and impressive. It all adds up to a two-part puzzle: Will it take an application distinct from the browser itself to bring back the notion of motion? And if so, does anybody still want to

Novell.

[Documentary Shows Web as Hall of Mirrors](#)



CULTURE
Today's Headlines

[Scorned Spammer Floods Net with Fake Threats](#)

[Documentary Shows Web as Hall of Mirrors](#)

[Short Films Get Long Attention](#)

[Beyond HOPE Hacks into Big Time](#)

[Street Cred: Freud as Fraud](#)

Novell.

surf?

This article appeared originally in [HotWired](#).

Find related stories from the Web's top news sites with [NewBot](#)



Feedback: Let us know how we're doing.

Tips: Have a story or tip for Wired News?
Send it.

DELIVERED

HOTWIRED

Copyright © 1993-97 Wired Ventures Inc. and affiliated companies.

All rights reserved.

The Seybold Report on Internet Publishing August, 1997

Alexa gets you where you want to go

Navigation service alleviates search engine roulette

It's no secret that search engines are now one of the weaker links on the Web. When Web users go to a search site, they usually expect plenty of duplicate URLs, irrelevant sites and links to pages that vanished weeks or even months ago.



Typical page at Alexa.com. The "where to go next" menu, along with the basic toolbar on the bottom, which gives the four most common navigation options, basic server info, and buttons for contacting the archive and online communication features.

Alexa Internet, a San Francisco-based start-up, is offering a product that may provide a real alternative to search engine roulette. The company's Alexa navigation service is, in equal parts, a search engine, a directory service à la Yahoo, an archive, and a collaborative filtering system. The resulting combination is a novel—and effective—tool for navigating the Web.

The Alexa service consists of a client application that appears as a toolbar at the bottom of a user's display. When the user goes to a Web site—a travel agency, for example—the client simultaneously sends the URL to the Alexa server, which suggests other Web sites that offer related content. Even if the initial link is lame and obscure (try entering "travel agency" in Alta Vista and see what you get!).

Alexa can point to higher quality sites—Travelocity, perhaps, or TravelWeb. If the user clicks through to a related but somewhat different site, such as a travel firm that deals specifically with cruises, Alexa will note the distinction and display updated links that match the user's interests.

Alexa works, first, by regularly crawling and archiving the entire public Web (about two terabytes of data) at the rate of around a terabyte per month. The service, which has been certified by the nonprofit TRUSTe (formerly known as e-Trust) initiative, honors the robots.txt protocol and will ignore sites at the owner's request. Alexa also regularly turns over its archives to the Internet Archive, a nonprofit organization dedicated to preserving a series of historical "snapshots" of the Web.

While Alexa archives the Web, the system uses several techniques to analyze sites. First, Alexa uses neural-net technology to perform a text analysis of sites, looking for pages and sites with similar characteristics. The system also analyzes link structures, detecting patterns that point to particularly common or popular sites. The idea here is to exploit the expertise built into the thousands of link lists and "favorite sites" pages, in almost every conceivable category, already available on the Web. In addition, the system makes creative use of public databases to find companies operating in the same business category—and that might have similar Web sites.

Alexa also includes data from the anonymized surfing habits of its users: which sites they visit, how deep they go, how long they linger and where they go next. Users can also explicitly state their opinions by voting on whether or not they found a given site helpful.

The system offers some other interesting features. Because Alexa archives Web content, users who click through to a "dead" page can receive an

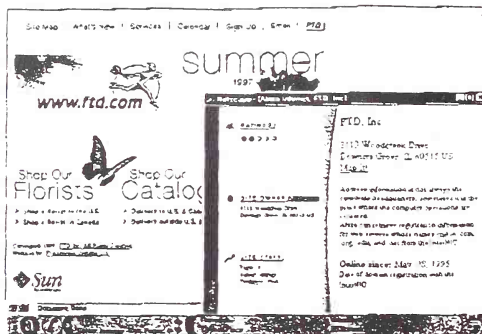
archived copy of the page instead of the usual "404—Not Found" error message. The toolbar can also provide information on who owns a site, how many pages are on the site, how many other Web sites point to the site, how frequently it is updated and how popular it is among other Alexa users. Finally, Alexa will show which other users are online and allow real-time communication with them.

Ads that stick. As always, the key question is "where's the money?" Alexa is using an ad-based business model, placing the ads on the toolbar and on an extended pulldown list of recommended sites. The ads are context-sensitive, so that users looking for that perfect travel site might see a Travelocity ad—even while they browse through competing sites.

While we wouldn't hazard a guess on whether this variant of Web advertising will fly when so many others have crashed and burned, it is certainly interesting. By keeping the ad on the Alexa toolbar, the service can keep an advertiser visible across a series of Web sites, something that even the biggest site-based ad networks can't promise yet. Alexa places ads only on designated areas of the toolbar, keeping plenty of free space for the system's objective choices.

Privacy issues might also be a concern here, but in this case Alexa has done everything right. The company guarantees that data from users will not be divulged to other firms. Alexa has also secured an endorsement from TRUSTe, a nonprofit consortium organized by the Electronic Frontier Foundation and CommerceNet, which monitors and certifies sites that adhere to its privacy guidelines. Users seem to appreciate the effort; according to Z Smith, the company's vp of engineering, more than 90% of the beta users have sent their e-mail addresses to use Alexa's instant messaging feature. Initially, the company

The Latest Word



Detailed view. The server information box includes the InterNIC summary on the owner's name, address and phone number; the site as ranked by Alexa users; and summaries of the server's speed, number of pages on the site, average freshness of pages and so forth.

had predicted that only 30% would do so.

We had an opportunity to see the Alexa system demonstrated, and then to try it ourselves. The results were impressive: Alexa was able to provide relevant links from some pretty obscure sites, and it did a good job of tailoring its recommendations as we surfed through sites with somewhat different types of content. We did encounter some bugs, such as the occasional off-topic recommendation list. Alexa is also rather slow at delivering pages out of its archive, a process that can take as long as 15 minutes. Part of this is inevitable, since the system keeps 90% of its data on tape robots. Part of it is also due to a rough caching model for what goes on the system's disk cache, where the other 10% of the data is available for higher speed access. Given the system's impressive showing at such an early stage in its development, however, we believe it will get even better as larger numbers of users come online and Alexa continues to refine its technology.

Alexa isn't going to replace traditional search engines, which will remain the best option for users who have a very specific query and a good idea of what they expect to get in return. Nor will it replace directory-

based services such as Yahoo, which can offer more complete lists of options organized in a distinct taxonomy. Alexa does, however, provide a convincing alternative for more general searches and context-based "where next?" queries, just the kinds of searches that make search engines so frustrating and painful to use. Alexa is, in short, a great idea; we hope the company can combine that idea with effective marketing and a sustainable business model.

Alexa is currently taking applications for its beta program, which will run through the summer. According to Smith, the company hopes to have the system ready for general use in a couple of months.

Matt McKenzie

Alexa Internet
Presidio of San Francisco
P.O. Box 23141
San Francisco, CA 94129-0141
Phone: (415) 561-6900
Fax: (415) 561-6798
www.alexai.com

SEPT.



SEPT.

Compilation of Alexa Press Coverage

As of September 3, 1997, Alexa has received coverage in the media:

Headline:	Publication:	Date:	Author:
Alexa's Theory of Relativity	PCWeek	8/18	Jim Rapoza
Cache This	Upside	8/18	Tish Williams
Download of the Week	C/Net TV - The Web	8/18	C/Net
Alexa Internet	Network News: Information for Business and Computer Applications	8/17	Alan Gatlin
FYI: Popularity Context	Star Tribune (Twin Cities/Minnesota)	8/4	New York Times
Scent of a Web site	Inman News Features	8/4	Inman News
The Web HotList	InfoWorld	8/4	Jim Battey
Internet boom just beginning, conference told	Reporter (Vacaville, CA)	8/3	San Francisco Chronicle
Internet industry 'not even close to peaking'	The Midland Reporter-Telegram (Midland, Tx)	8/2	Jon Swartz
Net Surf: Alexa's New Navigation Service	Wired News	8/1	Joey Anuff
Smart Searching at Internet World (Web Addict Section)	Web Review	8/1	Wes Thomas
Site Reviews: Alexa Internet	Media Central Digest	7/29	Sean Wolfe
Alexa: Merging Web Searching and Group Experiences	Internet Outlook (http://www.webreference.com)	7/28	Richard Wiggins
Alexa Keeps Surfers on Beaten Path	Inter@ctive Week (print)	7/28	Tom Steinert-Threlkeld
Alexa Searches While You	PC Week	7/28	Jim Kerstetter

Headline:	Publication:	Date:	Author:
Surf			
Blacks Say Internet Offers Equality	The Plain Dealer (Cleveland, OH)	7/28	From staff and wire reports
Officials See No End in Sight for the Popularity of the Internet	Post-Bulletin (Rochester, MN)	7/28	Jon Swartz (SF Chronicle)
New Service Tracks Web Use	Plainview Herald	7/28	New York Times Service
Revolution Brewing at Presidio: Alexa Internet hopes to change the way you use the Net	San Francisco Examiner	7/27	Zachary Coile
Net Growth Virtually Exploding	Austin American - Statesman (Austin, TX)	7/26	Jon Swartz (SF Chronicle)
The Internet Boom Is Just Beginning: Big business is the next frontier, attendees at Chicago trade show say	San Francisco Chronicle	7/25	Jon Swartz
Getting Even More Personal...	Interactive Services Report	7/25	No byline
E-commerce about to launch?	The Middletown Press (Middletown, CT)	7/25	Jon Swartz (SF Chronicle)
Outside The Box	San Jose Mercury News	7/24	Chris Nolan
As Go Surfers, So Goes Alexa	Wired News	7/24	Chris Oakes
Internet Industry Giants Snub Expo	The Desert Sun (Palm Springs, CA)	7/24	Jon Swartz (SF Chronicle)
2 Internet Giants Won't Show at Expo: Microsoft, Netscape are skipping Chicago	San Francisco Chronicle	7/23	Jon Swartz
New Ad-Backed Service to Monitor Context of Web Pages	Internet Advertising Report	7/23	Compiled by Beth Cox
Service Turns Page in Aiding Web Users	The Oakland Press (Pontiac, MI)	7/23	New York Times News Service
New Guide Offers Net Directions	Inter@ctive Week (online)	7/22	Tom Steinert-Threlkeld

Headline:	Publication:	Date:	Author:
New Search Tool Tracks User's Path Through Internet	The San Diego Union-Tribune	7/22	New York Times News Service
News.Scan: Today's Headlines from the Mainstream Press	PC Magazine Online	7/22	NewsWatch
Looking for a Second Commercial Hit, Brewster Kahle Launches a Navigation Service	Web Week	7/21	Margaret McKegney
Alexa Internet Lands at the Presidio	San Francisco Business Times	7/21	Steve Ginsberg
New Service Tracks Web Use	Nando.net (Reprint of NY Times Article)	7/21	John Markoff
New Service Tracks Web Use	New York Times	7/21	John Markoff
Alexa Internet: The Search as a Communal Effort	New York Times (CyberTimes Extra)	7/19	Laurie J. Flynn
Recent & Decent Web Tools	Family PC News Online	7/18	Michelle Megna
Alexa Browser Companion Makes Searches More Certain	PC World Online	7/18	Lisa Moskowitz
Re-engineering Surfing	San Jose Mercury News Online (link to TechWeb story)	7/18	Patricia Sullivan, Online Editor
Internet Historian Creates Tool to Inform Web Site Visitors	ZDNN-- The ZDNet News Channel	7/18	Renee Deger
Alexa Debuts Search Engine That Learns from Its Users	ZD Internet MegaSite Magazine	7/18	no byline
Startup Seeks To Synthesize Web	TechWeb	7/17	John Gartner
Alexa Internet Introduces Web Navigation that Learns from People	LA Times, BizWire Section	7/17	Press Release
Alexa Makes the Web History	C/Net	7/9	Alex Lash
Personal Access: Information abounds on the Internet, which means	Tucson Citizen	6/30	Leslie Miller

Headline:	Publication:	Date:	Author:
virtually anyone can find out virtually anything about someone else			

Microsoft Daily News 9/25/97	Click here to get the full story	Windows® Platform
Microsoft delivers Windows NT® 5.0 beta to more than 200,000 developers		

PCWORLD online	Join Now!	Index	Site Search	Help
----------------	-----------	-------	-------------	------

Alexa Searches So You Don't Have To

by Lisa Moskowitz

September 24, 1997

What if you didn't have to rely on keywords or search engines to pinpoint information on the Web? Say hello to Alexa 1.0, a free Web navigation service that works with your browser to guide you through the ever-growing maze of pages that is the Web.



Alexa is a gray toolbar that floats on your screen or neatly anchors beneath the browser window. When you land on a site, a pop-up box on the toolbar gives you a list of up to ten related sites you might be interested in visiting. For example, if you go to CNN.com, Alexa will suggest the MTV, New York Times, ESPN Sportszone, or Yahoo sites as possible next Web destinations. These "Where To Go Next?" listings are based on Alexa's searches for related content and by the surf patterns of other Internet users. Sometimes the suggestions are right on target. Other times they appear to be totally random.

In addition to the "Where To Go Next?" feature, Alexa also has a "Where Am I?" pop-up box that gives you data on a particular site. For instance, you can find out who owns the site, how many pages it has, its popularity among all Web users (based on the number of hits the page has received), how often its pages are updated, and how quickly they load. Also on the toolbar, you'll find online help and a link to the Alexa site. You can vote on the site you're currently visiting, send messages to other Alexa users, and retrieve pages from the Alexa Archive if you're looking for an older version of a page. A new feature in the final version is access to *Encyclopaedia Britannica* online. Abstracted entries are available, as are a *Webster's* dictionary and thesaurus.

Although Alexa is a useful tool that helps narrow your Web search, it does have its drawbacks. Toolbar icons aren't labeled and you have to put your mouse over them to reveal their functions. Less popular sites often don't have links, while the suggested links at some sites don't divulge the slightest clue as to what's on a page. Other links seem to work off of a site's server rather than its content. For example, I looked at a site about Chinese pug dogs on a server at the University of California at Davis. Instead of guiding me to other sites about pugs or dogs in general, Alexa zeroed in on the host server and suggested I visit environmental information sites next.

So it's not perfect, but Alexa is a pretty good tool for surfing the Web without having to jump back and forth between a site and a search engine. And even if the suggested links don't cough up the exact facts you went online to find, you'll end up exploring Web sites you might never have known existed. What better

[Download Alexa 1.0](#)

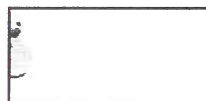
[Alexa Browser Companion Makes Searches More Certain](#)

[Alexa Searches So You Don't Have To](#)

[Powerful Pain Relief: Tips for Windows 95](#)

[More Daily News](#)

[Alexa Internet](#)



way to discover the true breadth of information available on the web.

Alexa works with Internet Explorer 3.0 or higher, Netscape Navigator 2.0 or higher, and Netscape Communicator 4.0. You can download it now by clicking on the link to your right.

Copyright © 1997 PC World Communications. All Rights Reserved. Use of this service is subject to the PC World Online Terms of Service Agreement.



The fastest, most secure way to transport
s over the Internet.

FileDrive EX Differential



Join now **FREE!**

MENU

[Front Door](#)

[The Net](#)

[Computing](#)

[Intranets](#)

[Business](#)

[CNET Radio](#)

[Perspectives](#)

[Newsmakers](#)

[Rumor Mill](#)

NEWS OPTIONS

[One Week View](#)

[Desktop News](#)

[News Alerts](#)

[Custom News](#)

[Advanced Search](#)

[Push](#)



RESOURCES

[Subscribe](#)

[Member Services](#)

[Contact Us](#)

[Help](#)

CNET SERVICES

[CNET.COM](#)

The Net

Alexa accessorizes browsers

By [Alex Lash](#)

September 22, 1997, 5:00 p.m. PT

[Alexa Internet](#) released today the first commercial version of its reference and navigation tool that plugs into a browser and provides background information on Web sites.

Once downloaded and installed, the Alexa software, which is free and appears as a toolbar at the bottom of a user's browser, monitors a user's surfing to display statistics for Web sites, suggest related sites, and offer quick links to online reference material.

With the glut of information on the Web, not to mention a glut of search engines and directories, Alexa's in-browser tool could come in handy for those who don't want to click through to another site for ancillary information.

Alexa sees its product as similar to that of [Firefly Network](#) and others that seek out users' potential preferences, as well as search engines and directories such as [Yahoo](#) and [Excite](#).

"The advantage of what we do compared to Yahoo, for example, is that we stay with you as a dashboard wherever you go," said Alexa spokeswoman Cynthia Lohr. "Search engines currently don't do that."

A button on the toolbar links the user to thumbnail guides of the [Encyclopaedia Britannica](#), Webster's Dictionary, and Webster's Thesaurus.

Alexa uses intelligent agents, also known as "bots" or "crawlers," to gather and display information about each Web site the user visits.

FREE newsletter

[Subscribe](#) [sample](#)

Latest Headlines display on desktop

The Net

- [ANS expands 56-kbps access](#)
- [VDOnet to expand repertoire](#)
- [Upgrading business rules for e-commerce](#)
- [Prodigy ad gets in users' faces](#)
- [UUNet streams into multicasting](#)

Computing

- [Grove tells of "multiple bifurcation"](#)
- [Compaq builds to order in Japan](#)
- [IBM backs clustered NT servers](#)
- [Grove soft on NCs, Net PCs](#)
- [IBM retires Deep Blue from chess](#)

Intranets

- [Developers to Gates: Adopt Java](#)
- [MS bets the company on NT 5.0](#)
- [Netscape dials up telcos](#)

Business

- [Avant loses ground in legal battle](#)
- [Oracle FY98 growth beating Q1 levels](#)
- [NEWS.COM Corrections](#)
- [Verity turmoil reflected in Q1 report](#)
- [3Com full steam ahead](#)
- [TI catches DSP wave](#)

[BUILDER.COM](#)
[BROWSERS.COM](#)
[GAMECENTER.COM](#)
[SEARCH.COM](#)
[DOWNLOAD.COM](#)
[SHAREWARE.COM](#)
[ACTIVEX.COM](#)
[MEDIADOME](#)
[SNAP! ONLINE](#)

MARKETPLACE

[CNET Store](#)
[How to advertise](#)
[CNET jobs](#)

The Alexa toolbar instantly displays the following information about each site: who has registered the site; the number of pages on the site; how many other sites point to this site; how frequently the site is updated; and the site's popularity.

Alexa also suggests relevant links to each site by combining its users' feedback with the navigation paths of previous visitors to the site. The company hopes to garner revenue by selling space in the links' windows to advertisers, similar to the method search engines such as Yahoo use to match banner ads to users' keyword searches. The company will not sell the recommended links themselves, a spokeswoman said.

Because Alexa is the commercial arm of the [Internet Archive](#), it taps that resource to display Web pages that are no longer live. Instead of fetching a "not found" error message, for example, the end user will go to the archived version of an old page. The nonprofit Internet Archive project seeks to store for posterity Web pages, Usenet discussions, and shareware--anything Net-related that will be lost to history. Alexa gives back by adding to the archive any new sites its users come across.

Alexa users who know each others' email addresses can check to see who else is online. Without switching to a separate email client, those online can send each other instant text messages, which pop up from the toolbar.

The software currently runs on Windows 95 and NT. The company says it will have Macintosh, Windows 3.1, and Unix versions in the unspecified future.

related news stories

- [IBM unveils intelligent agent](#) July 31, 1997
- [Alexa makes the Web history](#) July 9, 1997
- [Intersect crawls for copyrights](#) June 2, 1997
- [Agents come in from the cold](#) May 12, 1997

Miss a day?

- [One Week View](#)

[Back to Top](#)

Copyright © 1995-97 CNET, Inc. All rights reserved.

Sept. 17, 1997

USA TODAY

NO. 1 IN THE USA... FIRST IN DAILY READERS

Tracking the Web's beaten paths

His archive helps Brewster Kahle point the way through cyberspace

By Elizabeth Weise
USA TODAY

SAN FRANCISCO — When he really gets going about his ideas, such as archiving the entire Internet or mapping the Web by looking at the trails users leave, Brewster Kahle can sound like Robin Williams — if Williams had gone to MIT.

Kahle hunches down beside a Coke machine-size box full of computer tapes that together hold the equivalent of about 50,000 books. His hands leap into the air, his voice shifts octaves, and suddenly he's two supercomputers trying to find a speed at which they can talk.

"How's this for you?" "Faster, it's faster." "Fine, how's this?" "No noise on this end." "Great, I'll start shipping data now." "Fantastic. I'm ready!"

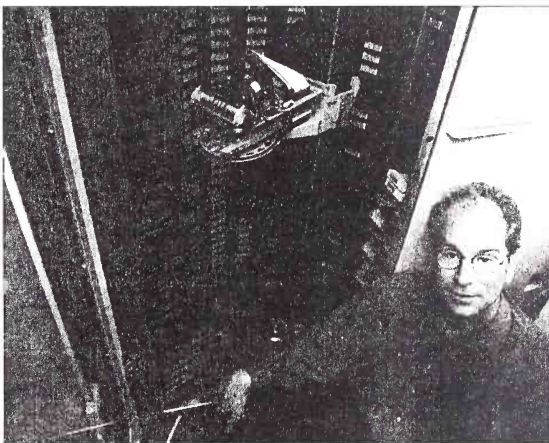
Described by Microsoft's chief technical officer Nathan Myhrvold as "a crazed lunatic, brilliant visionary and nice guy all rolled into one," Kahle has always thought big.

In the early 1980s, he was a scientist at a company called Thinking Machines, where he designed "massively parallel" computers, powerful machines made up of 100,000 small computers connected by a fast network. They broke huge problems into small bits that could be solved simultaneously.

Now Kahle, 36, is turning his expertise in the analysis of really big amounts of data to the problem of finding things on line. His new company, Alexa Internet, is a World Wide Web navigation service that gives users information about where they are and also recommends where to go next.

At his offices in a renovated Victorian general store in the former Presidio Army base here, Kahle talks a visitor through the service, his blue eyes glowing with pride under a cloud of curly blond hair.

Like a trained tracker, Alexa (<http://www.alexa.com>) helps guide users roaming the Web. Named for the lost library in ancient Alexandria, it creates a thin toolbar that shows a constant stream of information: who registered the site you're at, how often it's updated, how



By Julie Sogauer. AP. for USA TODAY

His Internet preservation act

Brewster Kahle founded the nonprofit Internet Archive in 1996 to record the cacophony of human voices being constantly created — and deleted — on line.

"The Net isn't 10 channels of TV. It's something fundamentally different and worth preserving." Whether a publisher hated your manuscript or democracy failed to take root in your country, "the Net is the answer," he says. The question was how to make sure historians would have access to it.

Using "Web crawling" programs, which copy everything they find and send it back to the archive's serv-

er in San Francisco, Kahle makes a full copy of the Web and Usenet discussion groups every six months.

The archive is stored on 500 tapes, each of which holds 50 gigabytes, the equivalent of about 50,000 books. Kahle estimates that a copy of the Web at its current size will take about 4 trillion bytes of data — 4 terabytes, or the equivalent of a good-size urban library. The Library of Congress holds about 20 terabytes of data.

Copies of all the tapes are also being stored in Seattle, where Kahle hopes eventually to build a think tank centered on the archives.

many pages it contains.

But it's as a guide that Alexa shines. Using anonymous data from Net traffic nodes, the service sees what paths others have taken and offers them to

users as a small pop-up list, ranked by which links were most heavily traveled.

"It's a sort of chatty navigator that in some metaphorical sense has talked with a lot of

Web visionary: Brewster Kahle, 36, with the collection of 500 computer tapes that hold a recent copy of the entire Web, the equivalent of 50,000 books.

people and can give advice about routes to take," says Jerry Michalski of Release 1.0, an industry newsletter.

And users never have to see the dreaded "Error 404 — page not found" message. If a Web page no longer exists, Alexa will find a recent copy in Kahle's Internet Archive and serve it up. Alexa must be downloaded to be used, it's ad supported and free to users.

Both Alexa and the archive spring from Kahle's fascination with libraries as founts of information. His interest dates to the late 1980s, when he developed the Wide Area Information Server (WAIS), a pioneering Net publishing system.

Although HTML won out as the primary publishing format, the method Kahle devised for indexing the Net became one of the most popular lookup tools of the time, so popular

that in 1995, America Online paid \$15 million for it. That gave Kahle the funds to pursue his dreams, one of which was preserving the digital past.

No one who knows Kahle is surprised that when he decided there should be an archive of the Internet, he just sat down and made one. He tends to pursue his passions wholeheartedly, whether they're technical problems to be solved or social networks to be knit.

Take the Thursday-night potluck dinners Kahle and his wife, Mary, have held for the past 10 years. Each includes a question, "What's the most interesting game you've ever played?" or "What's the strangest place you've ever slept?" Every guest is obliged to answer in the form of a story.

Kahle, who graduated from MIT with a degree in artificial intelligence and Eastern religions, delights in inviting newcomers to these meals. The couple's answer to the problem of meeting interesting people once they'd left college.

It's his engineer's "let's make something to fix this problem" attitude that has gotten Kahle so far. But sometimes his enthusiasm for ideas causes him to overlook practical considerations, observes one of the fathers of the Internet, Vinton Cerf, now at MCI.

Kahle has sidestepped the looming issues of copyright and privacy raised by copying Web pages without their creators' express permission. You can always take a page down, but if it's in the archive, anyone can still access it. Special coding can be included to prevent a page from being archived or indexed, but many people don't know this.

Kahle says that by not worrying about the details he's able to do things others think are impossible.

And that alone is enough to make Cerf, someone who knows a thing or two about big plans, respect Kahle. "He's contributed more than his fair share of interesting and innovative ideas."

Says Cerf, "I think Brewster is the kind of visionary who bears watching."

[Front page,](#)[News,](#)[Sports,](#)[Money,](#)[Life,](#)[Weather,](#)[Marketplace](#)The logo for USA TODAY, featuring the word "USA" in a large, bold, sans-serif font above the word "TODAY" in a smaller, bold, sans-serif font. The "T" in "TODAY" is stylized with a small graphic element.

Tech Report

Inside Tech

09/03/97- Updated 08:08 AM ET

- [FAQ, Tips and](#)

- [Tricks](#)

- [Web Column](#)

- [Hot Sites](#)

- [Tech News](#)

- [Tech Reviews](#)

- [Answer Desk](#)

- [Shareware Shelf](#)

- [Web Potholes](#)

- [Web Resources](#)

- [Tech Front](#)

Resources

- [Index](#)

- [Search](#)

- [Feedback](#)

- [What's hot](#)

- [About us](#)

- [Jobs at USA](#)

You are a database and access abounds

Today's typical Net surfer can get all kinds of information that would have taken a private eye to turn up in the past, says Tom Sterner of San Francisco.

He ought to know. Sterner is an investigator with Decision Strategies International, and he isn't at all surprised at heightened concern about privacy in the age of the Internet. "Generally people freak out about people knowing stuff about them," he says.

Sterner and most investigators use specialized databases, whose "prohibitive" expense keeps them from the reach of the layman.

But that doesn't mean you can't find a lot of the same material elsewhere.

"If information exists in one place, it exists in more than one place," says database expert Carole A. Lane, who will testify Tuesday at Federal Trade Commission hearings about privacy online.

She calls that rule "Lane's First Law of Information." And it has a corollary: "If information exists anywhere, no matter how carefully guarded, it exists somewhere else, where virtually anyone can gain access to it."

Lane makes her point in a new book titled *Naked in Cyberspace: How to Find Personal Information Online* (Pemberton Press, \$29.95).

The book has more than 500 pages detailing how information professionals dig up data, and where they turn when they hit a dead end. It holds a few privacy lessons too.

In addition to databases with identifying information and public records from criminal convictions to fishing licenses, Lane cites sources even private investigators might not think to look at, such as alumni newsletters posted online (for news about marriages, births, job changes) and even genealogical bulletin boards.

- Jobs at USA

TODAY

births, job changes) and even genealogical bulletin boards.

"Genealogists are very generous about sharing data, and oftentimes it includes who the family members are now, and where they are," she says.

Genealogy may sound innocuous, but it could expose information that compromises privacy: mother's maiden name, for example, is often a key to sensitive financial data.

"I don't remember ever seeing mothers' maiden names in any site I've seen," Sterner says. "I would consider that to be the equivalent of a PIN for information."

Among other personal information Lane cites online:

- Phone directories. Many are free on the Web; some have enhanced features that let you type in a phone number to get name and address or print out a map of the neighborhood. Some even include unlisted phone numbers. While these are not given out by phone companies, they can be gleaned from other sources, such as information from product warranty cards, which is often sold to direct mailers.
- Marketing lists. Companies on the Web advertise all flavors of lists that allow advertisers to target specific types of consumers, from occupations (Database America touts 250) to people who have moved recently, are "upscale grandparents" or own pets. "Even pharmacies sell databases, which shocked me," Lane says.
- Internet discussion logs. Web sites including Deja News (<http://www.dejanews.com>) keep searchable records of ongoing conversations on thousands of topics, from TV shows to business to quilting.

A student might post uncensored anecdotes or opinions, "not realizing that somebody years later, perhaps a potential employer, might learn they had done things they were embarrassed by," Lane says. "They might never tell you what they saw, but they might toss your resume before ever calling you."

It's now possible to specify that a posting in an Internet discussion or "newsgroup" not be archived, but not everyone knows this, and many others just don't bother.

- Home pages. Do-it-yourself personal pages proliferate on the Web; tools to create them and places to publish them are now free with many e-mail accounts. And people post all kinds of personal information "not thinking how it could be misused," Lane says.

Examples include everything from resumes with home address, phone and hobbies to pictures of people's kids.

"I wouldn't post my phone number and address at the local grocery store bulletin board," Lane says. "So why would I post it online on the whole world's bulletin board?"

Home pages can be changed or taken down, of course, but even now "snapshots" are being taken of what's on the Web at any given time.

Brewster Kahle, known for designing the first Net publishing system, WAIS, is now creating an Internet Archive (<http://www.archive.org>).

His digital robots crawl the Web making copies of every page so scholars and historians in the future will be able to study the early days of the Net in the same way they now use letters and diaries to study the past.

"The wonder of the Internet is it's changing rapidly," Kahle says. With the archive, "you can dial the time you want" and see what was there.

People seeking specific personal information wouldn't want to download the massive archive files, but Kahle plans this summer to launch a library-type service called **Alexa**, which will automatically direct browsing software to grab a recent copy of a page if it can't be accessed directly.

His goal is simply to make the Web "more reliable and less flaky," but Kahle obviously sees privacy implications.

So far, though, he says, "The spooky aspects haven't cropped up."

And if anybody doesn't want his page archived, Kahle doesn't argue. "We take it off."

By Leslie Miller, USA TODAY

- [Go to High-tech index](#)
 - [Go to Life Front](#)
 - [Go to Lifeline](#)
 - [Go to Money Front](#)
 - [Go to Moneyline](#)
 - [Go to News Front](#)
 - [Go to Nationline](#)
-

- [Go to Web Traveler](#)
-

Front page,

News,

Sports,

Money,

Life,

Weather,

Looking to upgrade

your development
bag of tricks?

"So what do I
need to protect
myself from
on the Internet
today?"

[Click Here!](#)


Welcome to
Dave

SOFTWARE ARCHIVE

Home
About
Advertise
Hot
Mirror
New
Search
Submit

- Home
- Audio
- Conferencing
- Connectivity
- E-mail
- Fax
- File Utilities
- Finger/Whois
- FTP
- Graphics
- Java
- Multimedia
- NewsGroups
- Programming
- Real Time Info/Push Media
- Telnet
- Video
- VRML
- Web Authoring
- Web Surfing
- Accelerators
- ActiveX Controls
- Bookmark Managers
- Browser Utilities**
 - Abuse Net
 - ActiveXCavator
 - Alexa
 - AutoBrowse
 - Collective Web
 - CopyURL
 - Cover Your Tracks
 - CyberSense Personal Start
 - Devil's StarterPage
 - Digigami Plugsy
 - Download Butler
 - DragonDROP
 - gReader PRO
 - Dropit
 - EasyFetcher
 - FAB Browser
 - Flying Toolbars
 - FormAgent
 - GetRight
 - HindSite
 - HotClipBoard 32-bit
 - ELinkMaster
 - IntelliLinks
 - JOY-IT
 - jpgWebReaper

Web Surfing Browser Utilities

[Previous Product](#) - [Next Product](#)

Alexa 1.0 Freeware

Alexa is a web navigation service that helps you surf the web smarter, faster, and easier. The service is displayed as a toolbar that works with your web browser (Internet Explorer or Netscape) to provide a continuous source of relevant recommendations of where to go next on the web.

With Alexa, you learn more about each site you visit, including who owns it, how many pages it has, how frequently it is updated, and its popularity among other Alexa users. Also, there is a list of links corresponding to each site that represent places you probably would like to visit. Alexa also has support for instant messages and searching Encyclopedia Britannica Online and Merriam-Webster's Dictionary and Thesaurus.



[Alexa Home Page](#)



Publisher: [Alexa Internet Vision](#)



[Download 32-bit version \(1.82M\)](#)



Platforms: Win 95, Win NT 4.0

[Previous Product](#) - [Next Product](#)

Dave

Home About Advertise Hot Mirror New Search Submit

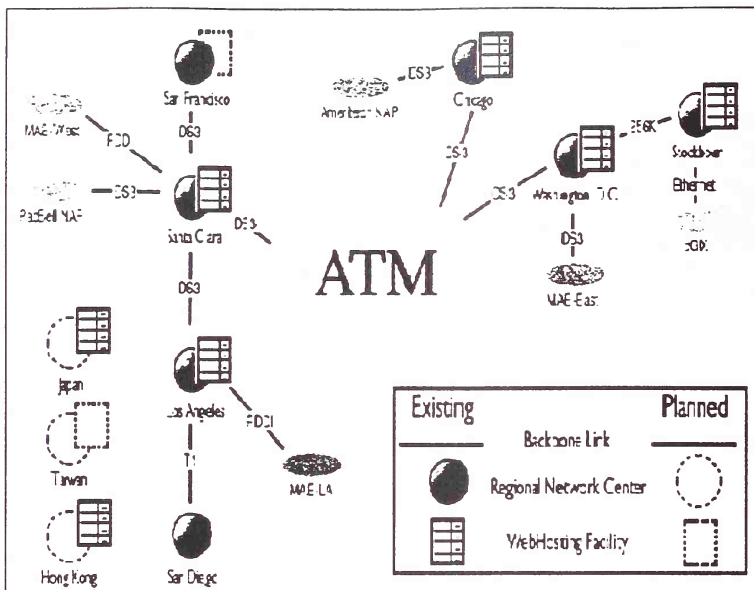


Figure 2. InterNex's Network

We operate with two licenses of Oracle 7.3.3. The databases are synchronized, but independent. Again, if one system is rendered unavailable, the second system will take over automatically.

TRUSTe runs three redundant Apache Web servers, but will soon move to Oracle's Application Server and implement Oracle's Payment Cartridge.

Alexa & WorldPages 'Point to' TRUSTe Licensees

Alexa Internet, a web navigation service, is working with TRUSTe on a service that will readily identify TRUSTe licensees to web users. Alexa works parallel to, but independently of, web browsers and is displayed as a toolbar on the bottom of the user's screen. In a nutshell, the bar knows where the user's browser is, describes "where you are", and provides input about "where to go next." A "where you are" pop-up box displays meta data for the site—including the TRUSTe logo of licensees.

Just as a recognizable trustmark can stimulate commerce, we believe TRUSTe can boost traffic to member web sites. How? Through an authorized directory—the "safe" place to browse. WorldPages, a leader in Internet database look-up services, is working with us to develop the directory. WorldPages will contribute web page design, a "million banner impressions" to drive traffic to the directory, promotion to all major search engines and sites, a "link to us" program, and home page links to TRUSTe.org. All TRUSTe licensees will be featured free via a searchable, alphabetical categorization system.

UMBC AgentNews Webletter

UMBC LAIT | AgentWeb | AgentNews | [Subscribe](#) | [Search](#) | 96 | 97: 1 2 3 4 5 6 7 8 ...

73. It is not a language's weakness but its strengths that control the gradient of its change: Alas, a language never escapes its embryonic sac.

-- Alan J. Perlis, "Epigrams in Programming", ACM SIGPLAN, September, 1982.

THIS ISSUE

Volume 2, Number 08
September 13, 1997
Baltimore, MD

<http://www.cs.umbc.edu/agents/agentnews/1997/08/>
Modified Saturday, 13-Sep-97 23:27:25 EDT

AGENT NEWS

Agents'98



Agents'98, the Second International Conference on Autonomous Agents, will be held Sunday May 10-Wednesday May 13, 1998 in Minneapolis/St Paul. Agents'98 is sponsored by ACM/SIGART and in cooperation with AAAI. Papers are due on October 1, 1997 with instructions available at <http://www.aaai.org/Conferences/AA98ss/submit.html>. For information about proposing tutorials, workshops, demonstrations and exhibits see the Agents'98 web page. Send email to agents-98-subscribe@cs.umbc.edu to be automatically added to a mailing list for occasional news about the conference.

Positions available

New agent-related positions were posted by USC/ISI, The University of Calgary, Department of Mechanical Engineering, Kinetic, Charles River Analytics, FCI Communications, Manchester Metropolitan University, ZUNO (UK), and IS Robotics

AGENT TECHNOLOGY

Jess 3.0

Ernest Friedman-Hill of Sandia has released Jess 3.0. This is a final release of Jess, a clone of the popular CLIPS expert system shell written entirely in Java. Jess 3.0 adds significant new features like multifields, multislots, and incremental reset, includes a number of bug fixes, a more flexible user function interface, *lots* of new built-in functions, and a general code cleanup and reorganization (a 'jess' package.) Jess 3.0 is compatible with all versions of Java starting with version 1.0.2. This (in particular) Java 1.1.3 compatible.

Alexa



Alexa Internet (founded in April 1996 by Brewster Kahle and Bruce Gilliat describes itself as "The company gathers, manages and analyzes multi-terabyte collections of information to make information access and navigation easy for everyone." Their first product is Alexa -- a client program (as opposed to a plug-in or proxy server) that has two components -- the toolbar and the contact system. As you browse the Web, the Alexa toolbar contacts their servers independently and separate from your Web browser. Alexa provides the following services:

- Where Am I? -- Alexa provides context for a site visited by offering instant background information (e.g., number of pages, frequency of updates, number of links into the site, popularity, etc.) so you can decide if it is worthwhile.
- What happened to my page? -- Alexa gives you automatic access to unavailable web pages through their archive, eliminating the problem of "404 Not Found" messages.
- Where should I go next? -- Wherever you travel online, Alexa is there with suggestions hot-linked to places you might want to visit next.
- Who else is online? -- Alexa tells you instantly which of your friends and colleagues are currently online, and lets you send an instant message so you can communicate with them in real-time.

The first two services are provided using their own archive of the Web which, as of July 1997, is in excess of five Terabytes. Alexa began collecting the Web in early 1996 and now has at least one snapshot of over 500K web sites. They adjust site visits to reflect their apparent speed of change and estimate that they have a new snapshot approximately every 60 days. The browsing suggestions are computing using collaborative filtering over the database of paths taken by Alexa users (with suitable privacy safeguards).

Search Engine Watch

Search Engine Watch is a nice collection of information on web search engines. Creator and maintainer Danny Sullivan, of Calafia Consulting, say "It contains all that "A Webmaster's Guide to Search Engines" had, plus it is more accessible to search engine users. The site remains grounded in the technical. How do these things work, or more specifically, how well do these things work. It also remains committed to tracking search engine news, important changes, and providing quality information about these important tools. The information from continuing studies, search engine help pages, articles, reviews, books, tips from others, feedback received directly from the various search engines, and interviews I conduct in the course of covering the search engine beat." They also run a mailing list devoted to search engines.

Now that parents
in Charlotte can see
school work online,



MERCURY CENTER | SAN JOSE MERCURY NEWS

TALK IS CHEAP

- [Back to Chris Nolan's Talk is Cheap page](#)
- [Visit the Talk is Cheap Library](#)

[Home](#)[NewsLibrary](#)[Index](#)[Feedback](#)

Silicon start-ups on the silver screen

Sept. 8, 1997

BY CHRIS NOLAN
Mercury News Staff Writer

SO you think ya oughta be in pictures, eh?

Well, you're not alone. Start-up veterans **Eric Swildens** and **Jay Capela** are making a movie about their -- and maybe even your -- lives. But unlike many of those who envision immortality in pictures, Swildens, the executive producer, and Capela, the producer and director, have the money and experience to make this film.

The film's working title is **Wild@Start** -- but this is no vaguely hallucinatory **David Lynch** tour of America. No, it's going to be a real-life, warp-speed tour of life in a Silicon Valley start-up. It's going to be a documentary about real companies and real people. In fact, said Swildens and Capela, making a movie is a lot like starting a company (To make the comparison complete, there's a [Web site](#)).

"There's a lot of excitement and terror," said Swildens. "All the time, it's 'We're doomed' or 'I'm the richest person in the world.'" Swildens, 30, isn't the richest person in the world or even the valley. He'll cautiously admit to being a millionaire as a result of selling his start-up, **Microline Software**, to **Neuron Data**. "The highs and the lows," he said, "were really high and low."

Director Capela, a film school veteran who worked for a while at **Swell Software**, a start-up since sold to **Adobe Systems**, said he's looking for the right balance of start-ups to feature in the movie. Capela said he's talked with a couple of high-profile companies like **Alexa Internet**, **EBay** and **Crossroute**. And the movie hopes to depict all the players in the start-up

process -- from venture capitalists to lawyers to established companies. Oh, and just like a start-up, Swildens and Capela are looking for investors, too.

Wild@Start is scheduled to start filming -- this is intended to be a theatrical release -- in about four weeks, said Capela.

WHERE DO YOU WANT TO GO TODAY? They had a contest to rename **CUC International** -- the software conglomerate that earlier this year purchased **Berkeley Systems** and **Books That Work** -- but no one won.

It's too bad, too, because the winner would have gotten dinner for two, anywhere -- and they meant anywhere -- in the world. The contest was open to the 15,000 employees of CUC, which is based in Stamford, Conn., although Books That Work and Berkeley are Bay Area companies. CUC got some 10,000 entries, according to spokeswoman Sue Tobin. "There were some really great names. We'd narrow it down to a few, but they'd be taken," she said.

The company couldn't take a name that's already registered with the U.S. Patent and Trademark Office. Names that were spelled different but sounded alike were out, too. Finally CUC jettisoned all the submissions, and resorted to an identity company.

There were rewards for good-faith efforts, though. Runners-up got American Express gift certificates worth \$150 to use for dinner, said Tobin. They just had to pay their own transportation to and from.

And CUC's new name? It's **Cendant**. The change is official later this month.

IMAGE OR REALITY: Two Silicon Valley companies have taken to television to sell their wares. But this isn't the usual pre-Christmas, back-to-school computer sales rush.

Instead, the ads from **Cybercash** and **CrossRoads Software** push products the average household will never want or need: digital commerce technology from Cybercash and integration software from CrossRoads.

The ads are running on networks with a rather particular audience -- CNN, CNN/fn and CNBC, whose up-to-the minute breaking news and stock tickers serve as the background video of preference for brokerages, investment banks and, we confess, newsrooms. These are the kinds of places that can dramatically affect a company's stock price.

"The market is largely investors or brokers and people who are connected to the financial community," said **Brian Rolfe**, a spokesman for Cybercash. He said the ads are also targeted at potential customers -- CEOs, CFOs and CIOs. "It's both. It's absolutely both," he said. "Who's going to get it the most, I'm curious to find out."

Not surprisingly, there's an agenda here: CrossRoads wants to go public. And Cybercash -- called Cybercrash by some investors because the price has fallen

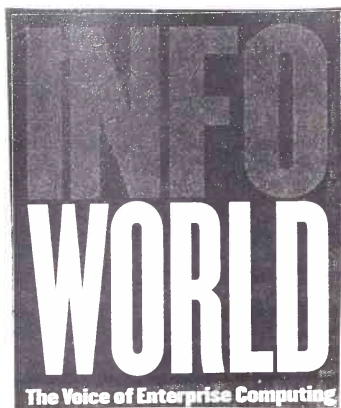
so far, so fast -- needs help getting its stock back to something like the \$40.50-a-share high of a year ago. It's now trading at about \$18 a share.

"You're right, it's not a very typical path," said CrossRoads CEO **Katrina Garnett**. But it's not expensive and it might work, she said. "I'd rather do it this way, than spend \$100,000 on a page in the Journal for one day."

Got a tip? A story? A good party? Pick up and phone and dial or log on and start typing. You can reach me at the San Jose Mercury News, 750 Ridder Park Drive, San Jose, Calif., 95190; 408)920-5490; or e-mail CNolan@sjmercury.com. Talk is Cheap appears Mondays and Thursdays.

Posted at 6:51 p.m. PDT Sunday, September 7, 1997

[Mercury Center Home](#) | [Index](#) | [News Library](#) | [Feedback](#) | ©1996-7 Mercury Center. The information you receive on-line from Mercury Center is protected by the copyright laws of the United States. The copyright laws prohibit any copying, redistributing, retransmitting, or repurposing of any copyright-protected material.



February 23, 1998

WINDOW MANAGER • BRIAN LIVINGSTON

Alexa provides additional ways to navigate the Web

Now a new, free service can give you vital information on how popular a Web site is and how many people liked and disliked their interactions with it. Even better, you can jump to related sites without having to resort to a keyword search, and you can retrieve old Web pages that have dis-



WINDOWS USERS ARE increasingly exploring the World Wide Web. Fewer than 20 percent of Windows PCs don't have Internet access, according to a recent Microsoft press briefing. But looking at a Web site and having confidence that the site is legitimate and trustworthy are two different things.

appeared from the Net.

Alexa is a relatively small freeware download (918KB in size), that works with Netscape Navigator 3.0 or later and Internet Explorer 3.0 or later on Windows 95 and Windows NT. (If you are using Windows 95, you need a version of the file Kernel32.dll dated later than Feb. 2, 1996. You can download the updated version from Alexa's site.)

When you install Alexa, a small toolbar appears at the bottom of your screen when your browser is active. While you're viewing a Web site, clicking the Where Am I button on Alexa's toolbar reveals a wealth of information.

Leading the display is a ranking of the number of hits the site receives based on an analysis of traffic across the Internet backbone.

Even more useful is the name, address, and phone number of the site's owner.

This information is retrieved from the domain registrar InterNIC, and includes the length of time the domain has been in existence. Because longevity alone is not a good indicator of an Internet business' reliability, Alexa collects votes from its users on whether they liked a site. Unfortunately, you can't learn why visitors voted as they did. But in one case, for example, a particular travel service developed a large number of "dislike" votes — perhaps a hint that you should consider a different service for your travel needs.

One of the most remarkable features of Alexa is that it can retrieve old Web pages that have moved or changed. When you see the message "404 Not Found" in your browser, you click the Archive of the Web button and Alexa will attempt to display the most recent archived version of that Web page.

This feature grew out of the Internet Archive project, an effort to preserve Web pages for posterity. Officials at Alexa's parent company, San Francisco-based Alexa Internet, say it holds three different "snapshots" of more than 500,000 Web sites that have been collected since early 1996. This information occupies 8 terabytes (8,000GB) of storage — equivalent to all of the data in all of the tapes in a typical video store.

Alexa uses the links between sites, as well as its analysis of how users jump from site to site, to determine other sites that may be relevant to the site you're currently viewing. Clicking Alexa's Where to Go Next button lists these suggestions. The links aren't always pertinent, but they're a start. Alexa supports itself with small advertisements in this box.

You don't even need Alexa to use some of its data. For example, set your browser to <http://widener.alexa.com/sitedata/yahoo.com> to see contact and rating information on the Yahoo search engine site. Change yahoo.com to any site you wish.

But the download is well worth it. Go to <http://www.alexa.com/download>. I would like to thank Danny Sullivan at <http://searchengine.com> for this suggestion.

Brian Livingston is the co-author of several best-selling Windows books, including the most recent Windows 95 Secrets (IDG Books). Send tips to brian_livingston@info-world.com. He regrets that he cannot answer individual questions.

Linux develops new way to search Web

How it works

The Washington Post

By RAJIV CHANDRASEKARAN
The Washington Post

SAN FRANCISCO — For more than a year, a team of computer industry entrepreneurs has been using a handful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quirky venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate — and potentially profitable — use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines — Web sites, with names such as Infoseek and Excite, that look for other Web sites containing that keyword and then generate a list of sites. That can produce a long list to wade through and the search engines often miss relevant sites because the keyword is slightly off target.

But the Internet Archive's service — called Alexa, after the ancient library at Alexandria, Egypt — doesn't rely on a word search.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from

<http://www.alexa.com>

uses several powerful computers to crunch through the

archived data — a process called data mining — and look for patterns within the Web, ultimately providing a few recommended strands to travel.

From United Parcel Service's Web site, for instance, Alexa points the user to sites for Federal Express, the U.S. Postal Service and DHL Worldwide Express. Starting from a Web page of catalog clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

"Sometimes it's spooky how helpful it can be," said Brewster Kahle, the archive's founder and the president of Alexa Internet Inc., a San Francisco start-up company that intends to offer the new navigation service for free to Web users.

This is not the first time someone has claimed to have cracked the code for making the Web easily searchable. And it's not clear whether Alexa, which is being tested by about 10,000 people now, would function well if millions of people tried to use it. But in its early form it's getting generally upbeat reviews in the computer industry.

"Search engines can drown you in documents," said Jerry Michalski, managing editor of Release 1.0, a New York-based computer industry newsletter.

Asking the Infoseek engine to find sites with the keyword "shipping," for example, generates more than 291,000 citations with sites for a British port and an exotic bird store at the top of the list.

"Alexa is a big step forward," Michalski said. "It's like having a huge associative memory that's looked at most of the Internet."

If you want information on shipping a disc, using "shipping" as a keyword for a conventional search engine will produce hundreds of thousands of sites, including UPS's Alexa technology, starting from the UPS site, will look for:

- Similar words and phrases at other sites. For instance, Alexa might flag "logivers the package" as similar to "tracking number" from the UPS site and direct you to FedEx's site, which also contains phrases close to that.
- Other Web sites that have linked to the UPS site and include information on other relevant home pages.

- Traffic patterns. Alexa tracks where users were before arriving at UPS and where they went afterward, integrating these patterns into its guidance system. If Federal Express were the most common site where users went next, it would likely appear as a suggested site to move to.

After processing the above information, Alexa suggests 10 sites. From the UPS site, Alexa recently steered users to:

- FedEx.
- DHL Worldwide Express.
- U.S. Postal Service.
- Buckingham, Wash., Transportation Department.
- B&O Medlogix, a Dutch container shipping company.

Alexa also provides a brief dossier for each suggested site, such as the company running the site, the speed of its computers, the number of pages in the site and a rating from other Alexa users.

Argus Leader Sioux Falls, SD September 7, 1997

Navigate Internet to track packages

If you want information on shipping a package, using "shipping" as a keyword for a conventional search engine will produce hundreds of thousands of sites, including UPS's. Alexa's technology, starting from the UPS site, will look for:

- Similar words and phrases at other sites. For instance, Alexa might flag "delivers the package" or "enter your tracking number" from the UPS site and direct you to FedEx's site, which also contains phrases close to that.

- Other Web sites that have "linked" to the UPS site and include information on other relevant home pages.

- Traffic patterns. Alexa tracks where users were before arriving at UPS and where they went afterward, integrating these patterns into its guidance system. If Federal Express were the most common site where users went next, it would likely appear as a suggested site to move to //

Internet searches improve

Alexa doesn't rely
on words alone

By RAJIV CHANDRASEKARAN
The Washington Post

SAN FRANCISCO - For more than a year, a team of computer industry entrepreneurs has been using a roomful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quixotic venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate - and potentially profitable - use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines - Web sites, with names such as Infoseek and Excite, that look for other Web sites containing that keyword and then generate a list of sites. That can produce a long list to wade through, and the search engines often miss relevant sites because the keyword is slightly off target.

But the Internet Archive's service - called Alexa, after the ancient library at Alexandria, Egypt - doesn't rely on a word search. Instead, Alexa, a software add-on for Web browsers that can be downloaded from <http://www.alexa.com>, uses several powerful computers to crunch through the archived data - a process called data-mining - and look for patterns within the Web, ultimately providing a few recommended strands to travel.

From United Parcel Service's Web site, for instance, Alexa points the user to sites for Federal Express, the U.S. Postal Service and DHL Worldwide Express. Starting at the Web page of catalog clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

"Sometimes it's spooky how helpful it can be," said Brewster Kahle, the archive's founder and the president of Alexa Internet Inc., a San Francisco start-up company that intends to offer the new navigation service for free to Web users.

New service makes it easier to surf the Web

BY RAJIV CHANDRASEKARAN
THE WASHINGTON POST

SAN FRANCISCO — For more than a year, a team of computer industry entrepreneurs has been using a roomful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quixotic venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate — and potentially profitable — use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines — Web sites with names such as Infoseek and Excite, that look for other Web sites containing that keyword and then generate a list of sites that can produce a long list to wade through, and the search engines often miss relevant sites because the keyword is slightly off target.

But the Internet Archive's service — called Alexa, after the ancient library at Alexandria, Egypt — doesn't rely on a word search.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from <http://www.alexa.com>, uses several powerful computers to crunch through the archived data — a process called data mining — and look for patterns within the Web, ultimately providing a few recommended stands to travel.

From United Parcel Service's Web site, for instance, Alexa points the user to sites for Fed-

Alexa analyzes keyword traffic

If you want information on shipping a package, using "shipping" as a keyword for a conventional search engine will produce hundreds of thousands of sites, including UPS's. Alexa's technology, starting from the UPS site, will look for:

- Similar words and phrases at other sites. For instance, Alexa might flag "delivers the package" or "enter your tracking number" from the UPS site and direct you to Fed Ex's site, which also contains phrases close to that.

- Other Web sites that have "linked" to the UPS site and include information on other relevant home pages.

- Traffic patterns Alexa tracks where users were before arriving at UPS and where they went afterward, in-

tegrating these patterns into its guidance system. If Federal Express were the most common site where users went next, it would likely appear as a suggested site to move to.

After processing the above information, Alexa suggests 10 sites. From the UPS site, Alexa recently steered users to:

- Fed Ex.
- DHL Worldwide Express.
- U.S. Postal Service.
- Bellingham, Wash. Transportation Department.

- P&O Nedlloyd, a Dutch container shipping company.

Alexa also provides a brief dossier for each suggested site, such as the company running the site, the speed of its computers, the number of pages in the site and a rating from other Alexa users.

eral Express, the U.S. Postal Service and DHL Worldwide Express. Starting at the Web page of catalog clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

"Sometimes it's spooky how helpful it can be," said Brewster Kahle, the archive's founder and the president of Alexa Internet Inc., a San Francisco start-up company that intends to offer the new navigation service for free to Web users.

This is not the first time someone has claimed to have cracked the code for making the Web easily searchable. And it's not clear whether Alexa, which is being tested by about 10,000 people now, would function well if mil-

lions of people tried to use it. But in its early form it's getting generally upbeat reviews in the computer industry.

"Search engines can drown you in documents," said Jerry Michalski, managing editor of Release 1.0, a New York-based computer industry newsletter.

Asking the Infoseek engine to find sites with the keyword "shipping" for example, generates more than 291,000 citations, with sites for a British port and an exotic bird store at the top of the list.

"Alexa is a big step forward," Michalski said. "It's like having a huge associative memory that's looked at most of the Internet."

Alameda Times - Star
Oakland, CA
September 8, 1997

Internet service mines mountain of data

THE WASHINGTON POST

SAN FRANCISCO — For more than a year, a team of computer industry entrepreneurs has been using a roomful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quixotic venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate — and potentially profitable — use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines — Web sites, with names such as Infoseek and Excite — that look for other Web sites containing that keyword and then generate a list of sites.

That can produce a long list to wade through, and the search engines often miss relevant sites because the keyword is slightly off target.

But the Internet Archive's service — called Alexa, after the ancient library at Alexandria, Egypt — doesn't rely on a word search.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from <http://www.alexa.com>, uses several powerful computers to crunch through the archived data — a process called data mining — and look for patterns within the Web, ultimately providing a few recommended strands to follow.

“*Sometimes it's spooky how helpful it can be.*”

Brewster Kahle,
president, Alexa Internet Inc.

From United Parcel Service's Web site, for instance, Alexa points the user to sites for Federal Express, the U.S. Postal Service and DHL Worldwide Express. Starting at the Web page of catalog clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

“Sometimes it's spooky how helpful it can be,” said Brewster Kahle, the archive's founder and the president of Alexa Internet Inc., a San Francisco start-up company that intends to offer the new navigation service free to Web users.

This is not the first time someone has claimed to have cracked the code for making the Web easily searchable. And it's not clear whether Alexa, which is being tested by about 10,000 people now, would function well if millions of people tried to use it. But in its early form, it's getting generally upbeat reviews.

“Search engines can drown you in documents,” said Jerry Michalski, managing editor of Release 1.0, a New York-based computer industry newsletter.

“Alexa is a big step forward,” Michalski said. “It's like having a huge associative memory that's looked at most of the Internet.”

How Alexa's searches work

THE WASHINGTON POST

If you want information on shipping a package, using “shipping” as a keyword for a conventional search engine will produce hundreds of thousands of sites.

Alexa's technology, starting from the UPS site, will look for:

► Similar words and phrases at other sites.

► Other Web sites that have “linked” to the UPS site and include information on other relevant home pages.

► Traffic patterns. Alexa tracks where users were before arriving at UPS and where they went afterward, integrating these patterns into its guidance system.

After processing the above information, Alexa suggests 10 sites.

Alexa also provides a brief dossier for each suggested site, such as the company running the site, the speed of its computers, the number of pages in the site and a rating from other Alexa users.

internet WORLD™

Q&A | Brewster Kahle

Pioneer of Net Searching Talks of Plan For a Directory

By Whit Andrews

BREWSTER KAHLE'S WAIS INC. promoted the idea that "search" was as big a part of interactivity as "click" in the early 1990s, a notion that still is gaining currency as search blanks come off back pages and onto the toolbar.

Kahle's new company—he sold WAIS to America Online in 1995 for \$15 million—is Alexa, which aims to promote something even more visceral, along the lines of "Show me where I am and where I should go." The earliest version of this technology was a sizable download application; it's been pared, but the ultimate goal is to create a skinny Web service with no technology barrier between a user's wanting it and his using it.

Whether Alexa is an idea that will aid newbies, its target market, in penetrating the Internet's mysteries is too soon to say. But Kahle has walked this path before, and it's interesting what he sees as being very different this time out.

Internet World: What has changed in terms of getting in the public eye since you got in a smaller public's eye with WAIS?

Brewster Kahle: Back in the WAIS days, it was a concept sell, it was a vision thing. It was, "Everybody's going to become a publisher, it's going to be on the Internet, and they're going to use WAIS." The way you sold that back then was top-down. So [Vice President] Al Gore was a proponent of WAIS. I went category by

category and said, "Who's the top publisher in that field?" and I shot for it. If you could sell those, then the rest would follow.

That's completely different from what we're seeing now. Where Alexa has got the sort of coolness that lands us on NPR and that type of programming—and that helps in the public relations, to have a big concept associated with you—consumers couldn't care less, really. If it's not something they can use right then, and that works really well, and that doesn't crash, it's not in their sphere. They're a harder group to sell, in fact, than Al Gore and the head of The New York Times.

IW: Do you think the ethic has changed in terms of must-not-crash?

Kahle: We've moved from a technology demonstration phase to "Let's have it be useful in what I'm trying to do." People are starting to say that the answer is out on the Net someplace, but you can't find it. That change is recent, in that we've hit a critical mass, and people are now applying this technology to their everyday problems. We in the Internet are not there yet.

IW: In both cases, WAIS and Alexa, you're promoting an intellectual experience that to some degree is beyond imagination, except that of the more visionary people. You suc-

ceeded with WAIS in being part of this new world. We all know what "search" means now. How do you cross that barrier and gain critical mass with Alexa?

Kahle: Probably analogy. What we're trying to do with Alexa is be an onscreen directory, which is close to what we have seen on the 500-channel satellites. My father, who hates the Web because it's so disorganized, has this satellite system, and one of the things he wanted to show me the most was the directory. It wasn't just "Channel 3 is CBS." It was, "This is CBS, and here's what's showing right now." So we call it an onscreen directory, and that really has worked for a lot of the newbies.

IW: I know that money can be a major issue for a concept play. Are you feeling pressure from investors, or customers, who want to see you making money?

Kahle: Yes, people are going to want to make sure you're going to make money. And it's not just customers. It's employees, and all of these sorts of important people. I find it's helpful that I've been part of a concept that went from "You've got to be crazy" to "Of course!" once before. The thing that helps now is all the IPOs.



Brewster Kahle

IW: Do you find the environment for getting developers and strategists different from 1992?

Kahle: It's so much easier now.

IW: Why is that? Usually I hear people say, "I can't hire anybody."

Kahle: You compared it to 1992. Around 1995, we saw the start of the flood of talented people. 1995, 1996, and 1997, we've seen an immigration into California, and lot of people trying to make their fortune.

search: AGENTS, SEARCH SERVICES

BUSINESS MONDAY

YOUR WEEK AHEAD

Internet Archive stands and delivers

■ **NOLAN**

from Page 1E

Olson Stephens & Co. threw a party at the San Francisco Museum of Modern Art featuring open sushi bars, bowls made of chocolate overflowing with golf-ball-size strawberries

and some of the tenderest lamb chops ever grilled. Those days — sigh — are over and Rob-

bio's new partner, Bank of America, is getting the blame, Wednesday's party at Chalk-



CHRIS NOLAN

TALK IS CHEAP

port partner in San Francisco's Kincaid Center, was sparsely attended. And last Monday's event at Pier 39's Under Water World — in the weathered we've been having that's some invitation — was less than happy, according to one bank executive.

"We aren't going to go to Chalkers," sniffed one of Robertson's more senior bankers. "Investors aren't, either." His excuse for not attending the party? What else? Bubbles.

No, the public party to attend during the BancAmerica Robertson Stephens Technology 98 confer-

ence was 2Bridge's launch fete at the Old Federal Reserve Bank Building in San Francisco. That

small but crowded party, which set the company back about \$60,000 and a lot of favors, featured a raw oyster bar, pâté and cute little roast beef sandwiches complete with fancy whole grain mustard.

As befits any event where martinis were served and swing bands played, some guests got gussed up. The most changed-by-his-attire award goes to 2Bridge's venture capitalist, Stewart Alsop, one of the evening's hosts. Just by putting on a tux and evening vest, Alsop, who is a member of one of journalism's more established families, proved that you can take the boy out of the patrician East Coast society but it's much harder to take patrician East Coast society out of the boy.

"It was our coming-out party," said 2Bridge PR director Cathy Brooks. She means that in the patrician East Coast sense. "More of a debutante ball, perhaps," she clarifies.

So are they jealous over at Robertson? Well, no, not really. First of all, there's something of a family connection, 2Bridge managing director Charlie Loucks' dad, Vernon, and Sandy Robertson are childhood buddies. And Robertson, founder of the bank that bears his name, said this year's reduction in festivities around the technology

conference only looks like a cut-back.

Instead of public events, Robertson said he and other members of the firm were hosting smaller dinners. He held three last week: Monday for 75 guests, Tuesday for 120 and Wednesday for another 75. The guests are mostly CEOs and management teams at companies about to go public as well as the bankers — from Robertson — who will take them there.

"If you give a big party, it's not as personal," said Robertson. Not only that, it can give other banks an edge since people attending a conference don't always feel compelled to spend their days and their nights with their hosts.

"You have attrition of people going out with our competitors," said Robertson. "We do it, too."

ON A CHAIN GANG: How hard do they work at Intel? This list, described by one Intel employee as "a humorous exaggeration of Intel's daily grind, sent and distributed by people who put in the long hours & reap the good bonuses" has been making the rounds inside and outside the Santa Clara-based chip manufacturer.

It's pretty funny. But don't think they have too much of a sense of humor down at HQ. The author declined to be identified, saying he was not interested in jeopardizing

his chances of being promoted. The journal, of course, owes a debt to David Letterman.

Top 10 Signs You Work for Intel:

10. You lecture the neighborhood kids selling lemonade on ways to improve their manufacturing process.

9. You get all excited when it's Saturday so you can wear sweats to work.

8. You refer to the tomatoes grown in your garden as "delivery address."

7. You find you really need PowerPoint to explain what you do for a living.

6. You typically eat out of vending machines as well as at the most expensive restaurant in town within the same week.

5. You think that "progressing an action plan" and "calendarizing a project" are acceptable English phrases.

4. You know the people at airport hotels better than your next-door neighbors.

3. You ask your friends to "think out of the box" when making Friday night plans.

2. You think Einstein would have been more effective had he put his ideas into a matrix.

And the No. 1 sign you work at Intel?

1. You think a "half day" means leaving at 5:00 p.m.

September 4, 1997

The Washington Post

BUSINESS

Seeing the Sites On a Custom Tour

New Internet Search Tool Takes Selective Approach

By Rajiv Chandrasekaran
Washington Post Staff Writer

SAN FRANCISCO
For more than a year, a team of computer industry entrepreneurs has been using a roomful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quixotic venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate—and potentially profitable—use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines—Web sites, with names such as InfoSeek and Excite, that look for other Web sites containing that keyword and then generate a list of sites. That can produce a long list to wade through,

and the search engines often miss relevant sites because the keyword is slightly off target.

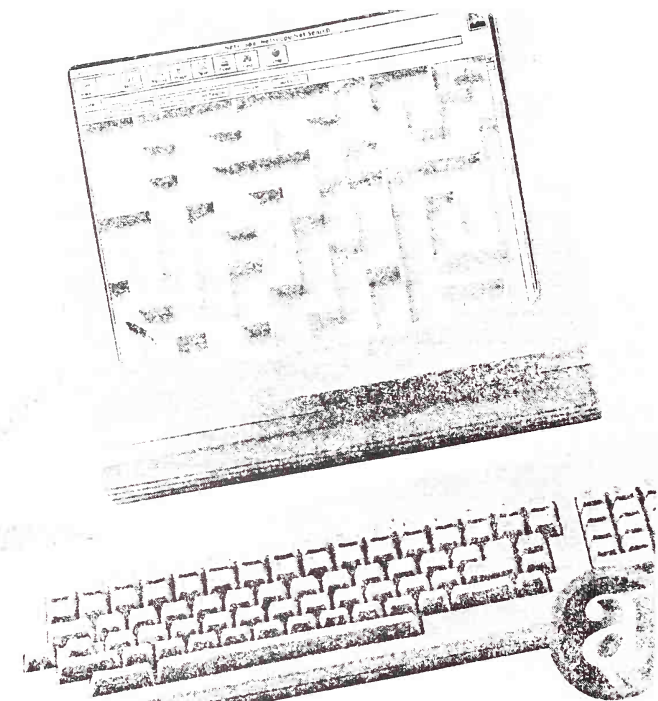
But the Internet Archive's service—called Alexa, after the ancient library at Alexandria, Egypt—doesn't rely on a word search.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from <http://www.alexa.com>, uses several powerful computers to crunch through the archived data—a process called data mining—and look for patterns within the Web, ultimately providing a few recommended strands to travel.

From United Parcel Service's Web site, for instance, Alexa points the user to sites for Federal Express, the U.S. Postal Service, and DHL Worldwide Express. Starting at the Web page of catalogue clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

"Sometimes it's spooky how helpful it can be," said Brewster Kahle, the archive's founder and the president of

See ARCHIVE, E8, Col. 1



NAVIGATING THE WEB

If you want information on shipping a package, using "shipping" as a keyword for a conventional search engine will produce hundreds of thousands of sites, including UPS's. Alexa's technology, starting from the UPS site, will look for:

- **Similar words and phrases at other sites.** For instance, Alexa might tag "de" as the "package" or "enter your tracking number" from the UPS site and direct you to FedEx's site, which also contains phrases close to that.
- **Other Web sites that have "linked" to the UPS site** and include information on other relevant home pages.
- **Traffic patterns.** Alexa tracks where users were before arriving at UPS and where they went afterward, integrating these patterns into its guidance system. If Federal Express were the most common site where users went next, it would likely appear as a suggested site to move to.

After processing the above information, Alexa suggests 10 sites. From the UPS site, Alexa recently stored users to:

- FedEx
- DHL Worldwide Express
- U.S. Postal Service
- Bellingham, Wash., Transportation Department
- P&O Nedlloyd, a Dutch container shipping company

Alexa also provides a brief dossier for each suggested site, such as the company running the site, the speed of its computers, the number of pages in the site and a rating from other Alexa users.

Chronicle Tribune
Marion, IN
September 19, 1997

Kahle is the man to watch

Gannett News Service

SAN FRANCISCO ~ When he really gets going about his ideas, such as archiving the entire Internet or mapping the Web by looking at the trails users leave, Brewster Kahle can sound like Robin Williams — if Williams had gone to MIT.

Kahle hunches down beside a Coke machine-size box full of computer tapes that together hold the equivalent of about 50,000 books. His hands leap into the air, his voice shifts octaves and suddenly he's two supercomputers trying to find a speed at which they can talk.

"How's this for you?" Faster, lots faster. "Fine, how's this?" No noise on this end. "Great, I'll start shipping data now." Fantastic. I'm ready!"

Described by Microsoft's Chief Technical Officer Nathan Myhrvold as a crazed lunatic, brilliant visionary and nice guy all rolled into one, Kahle has always thought big.

In the early 1980s, he was a scientist at a company called Thinking Machines, where he designed massively parallel computers, powerful machines made up of 100,000 small computers connected by a fast network. They broke huge problems into small bits that could be solved simultaneously.

Now Kahle, 36, is turning his expertise in the analysis of really big amounts of data to the problem of finding things on line. His new company, Alexa Internet, is a World Wide Web navigation service that gives users information about where they are and also recommends where to go next.

At his offices in a renovated Victorian general store in the former Presidio Army base here, Kahle talks a visitor through the service, his blue eyes glowing with pride under a cloud of curly blond hair.

Like a trained tracker, Alexa

(<http://www.alexa.com>) helps guide users roaming the Web. Named for the lost library in ancient Alexandria, it creates a thin toolbar that shows a constant stream of information: who registered the site you're at, how often it's updated, how many pages it contains.

But it's as a guide that Alexa shines. Using anonymous data from Net traffic nodes, the service sees what paths others have taken and offers them to users as a small pop-up list, ranked by which links were most heavily traveled.

"It's a sort of chatty navigator that in some metaphorical sense has talked with a lot of people and can give advice about routes to take," says Jerry Michalski of Release 1.0, an industry newsletter.

And users never have to see the dreaded "Error 404 — page not found" message. If a Web page no longer exists, Alexa will find a recent copy in Kahle's Internet Archive and serve it up. Alexa must be downloaded to be used; it's ad-supported and free to users.

Both Alexa and the archive spring from Kahle's fascination with libraries as founts of information. His interest dates to the late 1980s, when he developed the Wide Area Information Server, a pioneering Net publishing system.

Although HTML won out as the primary publishing format, the method Kahle devised for indexing the Net became one of the most popular lookup tools of the time, so popular that in 1995, America Online paid \$15 million for it. That gave Kahle the funds to pursue his dreams, one of which was preserving the digital past.

No one who knows Kahle is surprised that when he decided there should be an archive of the Internet, he just sat down and made one. He tends to pursue his passions wholeheartedly, whether they're techni-

cal problems to be solved or social networks to be knit.

Take the Thursday-night potluck dinners Kahle and his wife, Mary, have held for the past 10 years. Each includes a question, "What's the most interesting game you've ever played?" or "What's the strangest place you've ever slept?" Every guest is obliged to answer in the form of a story.

Kahle, who graduated from MIT with a degree in artificial intelligence and Eastern religions, delights in inviting newcomers to these meals, the couple's answer to the problem of meeting interesting people once they'd left college.

It's his engineer's let's-make-something-to-fix-this-problem attitude that has gotten Kahle so far. But sometimes his enthusiasm for ideas causes him to overlook practical considerations, observes one of the fathers of the Internet, Vinton Cerf, now at MCI.

Kahle has sidestepped the looming issues of copyright and privacy raised by copying Web pages without their creators' express permission. You can always take a page down, but if it's in the archive, anyone can still access it. Special coding can be included to prevent a page from being archived or indexed, but many people don't know this.

Kahle says that by not worrying about the details he's able to do things others think are impossible.

And that alone is enough to make Cerf, someone who knows a thing or two about big plans, respect Kahle. "He's contributed more than his fair share of interesting and innovative ideas."

Says Cerf, "I think Brewster is the kind of visionary who bears watching."

From a Mass of Web Sites, a New Road Map

ARCHIVE, From F1

Alexa Internet Inc., a San Francisco startup company that intends to offer the new navigation service for free to Web users.

This is not the first time someone has claimed to have cracked the code for making the Web easily searchable. And it's not clear whether Alexa, which is being tested by about 10,000 people now, would function well if millions of people tried to use it. But in its early form it's getting generally upbeat reviews in the computer industry.

"Search engines can drown you in documents," said Jerry Michalski, managing editor of Release 10, a New York-based computer industry newsletter.

Asking the Infoseek engine to find sites with the keyword "shipping," for example, generates more than 294,000 citations, with sites for a British port and an exotic bird store at the top of the list.

"Alexa is a big step forward," Michalski said. "It's like having a huge associative memory that's looked at most of the Internet."

Alexa works largely by looking for Web sites with similar content, then linking them together. A user starts with one Web site, and then the data-mining computers search for other pages that have several similar words and phrases. The computers suggest sites that are most closely related.

People who create Web pages also include links to other pages, but those links are often limited—Company A doesn't want people visiting its page to go to the page of its competitor Company B, and so it includes no link. Alexa, however, tries to function as an objective source as to what pages should be tied together.

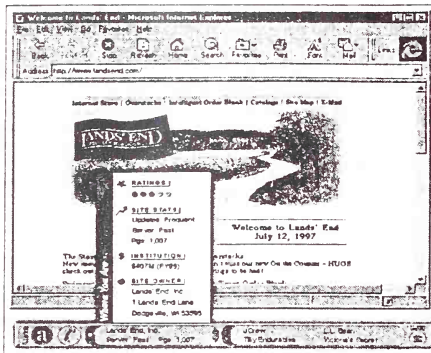
The machines also troll through the maze of connected pages that make up the Web, hunting for the most relevant sites that have links to the user's starting page.

Those suggestions, offered in a small box at the bottom of a user's screen, are shaped by the surfing patterns of other Alexa users, Kahle said. The sites they choose to jump to are tracked, those patterns will influence the 10 sites presented in the box.

For example, if many Alexa users at the UPS Web site don't jump to the FedEx site, it could be removed from the boxed list, even though the two sites are very common, Kahle said.

At the same time, if Alexa users at the UPS site tend to move to the AFL-CIO's Web page, that site could be suggested to future Alexa users visiting the UPS site, even though the two aren't associated in any way beyond the AFL-CIO's support for Teamsters in their recent strike against UPS.

• like asking somebody if they've



Under a "Where You Are" box, Alexa provides readily accessible information on the site you're visiting, such as company revenue and a rating of the site.

read any good books lately," Kahle said. He said that because of privacy concerns, Alexa's computers will not keep records of users' names as they jump from site to site.

Some early users question whether the service is detailed enough. Although it will suggest whole Web sites, it can't suggest specific Web pages that are similar to the page a user has on his screen.

"The database is a little immature," said Tiernan Ray, an analyst with Technology Partners in New York. "It can't give you much information about a specific page buried several layers into a site."

And its suggestions, Michalski said, "might take you to something related, but very random."

From the White House Web site, for example, Alexa suggests not just the House of Representatives, the FBI and the Library of Congress, but also PC Magazine and a service to look up stock quotes.

"It's not always perfect," Kahle conceded. "But we think it's a big improvement from the other ways of searching the Net today."

Part of Alexa's expected appeal, Michalski said, is its usefulness both in finding specific information and in simply cruising the Web.

"You get this list of sites, and some of them are going to be right on the money. And you can just hop to them without having to go back to your search engine," he said.

In addition to providing "where to go next" advice, the Alexa service can tap into another facet of the archive—its snapshots of the Web's past. Web pages of the Heaven's Gate cult, for instance,

are no longer actively offered on the Internet but can be culled from the archive here.

The service also tells users basic facts about the site they're visiting, including which company runs the site, how fast its computers are and how many Web pages the site contains.

"It's very useful information," Ray said. "It's something that most Web surfers have no idea about today."

Alexa, which has received \$5 million in funding from Etole, the Swiss investment company that owns Encyclopaedia Britannica Inc., plans to make money from the service through advertising, which will appear in the box at the bottom of the screen.

Kahle, who started Alexa with former Dow Jones & Co. executive Bill Dunn, eventually plans to merge some of the encyclopedia's content with Alexa's. Part of the profit Kahle hopes Alexa will generate could be used to expand and maintain the archive.

The service already has about 10,000 trial users and has the capacity to handle 1 million users, Kahle said.

Alexa also will let its users vote, with one mouse click, on whether they like a particular site. Those ratings will be offered to users.

"The Web allowed everyone with a computer to become a publisher," Kahle said. "Alexa allows them to become editors."

FOR MORE INFORMATION

To tour the Alexa site and preview its search method, click on the above symbol on the front page of The Post's Web site at <http://www.washingtonpost.com>

WIRED NEWS



✕ Kiss That 404 Goodbye

by Chris Jones

3:09pm 22 Sep 97 PDT What does one do with an archive of the entire public content of the Web? A new navigation service debuted today that aims to answer that question by giving users a free toolbar with which to explore a more comprehensive yet focused rendering of the Internet.

The [Alexa](#) Internet navigation service is a searching and reference tool that helps users find relevant Web sites and research-related information via the Britannica encyclopedia and Webster's dictionary and thesaurus. The Alexa tool recommends sites based on previous usage patterns, hyperlink information, and data collections from other sources. This metadata is analyzed by Alexa's data mining and collaborative filtering engines, but rather than returning hundreds, or even thousands, of URLs - à la AltaVista - the tool distills 10 sites it deems most relevant.

"We're not competing with directory sites like Yahoo. We try to keep you on track by giving more precise recommendations based on what other people say. It's more precise than other search engines, and has a much higher coverage of the Web," said Brewster Kahle, co-founder of Alexa.

Once downloaded, the Alexa navigation bar launches in sync with a browser, and consists of a handful of icons encapsulated in a slim, rectangular box. The toolbar supplies information about a site - who it's registered to, how many pages it contains, how frequently it's updated - and recommends other sites that include similar information. The tool also delivers ratings - XXX, for instance - that reflect the content of the site, and includes TRUSTe ratings that rate the privacy and data-collection policies of a site.

Kahle has in the past few years been involved with the nonprofit Internet Archive project, which aims to archive as many files as possible - from the Web and FTP sites, Usenet, and other sources. With about 640,000 sites, and 100 million pages, Kahle said the system has 7 terabytes of information stored. But with the number of sites on the Net doubling every six months - and existing sites adding new depths and layers all the time - keeping pace becomes a daunting task.

One advantage of having a full archive of the Web is that if a user gets the dreaded "404 -

TECHNOLOGY
Today's Headlines

[Net Telephony - Without the Net](#)

[Virtual Reality Goes Nuts and Bolts](#)

[XML Ushers in Structured Web Searches](#)

[Street Cred: Petite Plume](#)



TECHNOLOGY

GET STOCK QUOTE

GO

SEARCH



W



enter email

GO

The Times - News
Twin Falls, ID
September 21, 1997

Alexa service offers new way to navigate the Web

By Rajiv Chandrasekaran
The Washington Post

SAN FRANCISCO — For more than a year, a team of computer industry entrepreneurs has been using a roomful of computers to copy every World Wide Web page they can find.

Their massive digital library originally was viewed as a quixotic venture to assemble an Internet time capsule of the global computer network's early days.

But the electronic librarians appear to have found a more immediate — and potentially profitable — use for their 2 trillion-character pile of data: a new and precise way for ordinary computer users to navigate the sprawling Web.

To find what they're looking for on the Web, most computer users plug "keywords" into search engines — Web sites with names such as Infoseek and Excite, that look for other Web sites containing that keyword and then generate a list of sites. That can produce a long list of waste through, and the search engines often miss relevant sites because the keyword is slightly off target.

But the Internet Archive's service —

Navigating the Web with Alexa

If you want information about a pack page, using "Shipping" as a keyword for a conventional search engine will produce hundreds of thousands of sites, including UPS sites, which will look for:

- **United World and Phages** at other sites. For instance, Alexa might flag "delivers the package" or "enter your tracking number" from the UPS site and direct you to Fed Ex's site, which also contains phrases close to that.
- **Other Web sites** that have "linked" to the UPS site and include information on other relevant home pages.
- **Traffic patterns.** Alexa tracks where users were before arriving at UPS and where they

went afterward, integrating these patterns into its guidance system. If Federal Express' were the most common site where users went next, it would likely appear as a suggested site to move to.

After processing the above information, Alexa suggests 10 sites. From the UPS site, Alexa recently steered users to:

- DHL Worldwide Express.
- U.S. Postal Service.
- Bollingham, Wash., Transportation Department.
- P&O Nedlloyd, a Dutch container shipping company.

Alexa also provides a brief dossier for each suggested site, such as the company running the site, the speed of its computers, the number of pages in the site and a rating from other Alexa users.

called Alexa, after the ancient library at Alexandria, Egypt — doesn't rely on a word search.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from <http://www.alexa.com>, uses several powerful computers to crunch through the archived data — a process called

catalog clothing retailer Lands' End, sites for J. Crew, L.L. Bean and Victoria's Secret are suggested.

"Sometimes it's spooky how helpful it can be," said Brewster Kahle, the archive's founder and the president of Alexa Internet Inc., a San Francisco start-up company that intends to offer the new navigation service for free to Web users.

This is not the first time someone has claimed to have cracked the code for making the Web easily searchable. And it's not clear whether Alexa, which is being tested by about 10,000 people now, would function well if millions of people tried to use it. But in its early form, it's getting generally upbraided in the computer industry.

"Search engines can drown you in documents," said Jerry Michalski, managing editor of Release 10, a New York-based computer industry newsletter.

Asking the Infoseek engine to find sites with the keyword "shipping," for example, generates more than 291,000 citations, with sites for a British port and an exotic bird store at the top of the list.

"Alexa is a big step forward," Michalski said. "It's like having a huge associative memory that's looked at most of the Internet."

OCT.

OCT.

Compilation of Alexa Press Coverage

October media coverage:

Headline:	Publication:	Date:	Author:
fyi	Windows Magazine	11/97	Staff
Site Reviews: Go There Find That	Yahoo! Internet Life	11/97	Staff
Alexa Party	Upside	11/97	Tish Williams
Internet Search Engine	Newscenter Four Daybreak KRON-TV San Francisco	10/21	Stan Bunger
Internet Information	King Five Morning News KING-TV Seattle, WA	10/19	Mike Wendland
Net Surfin'	Weekend Sunrise WJAR-TV Providence	10/18	Mike Wendland
Internet	Today in St. Louis at Six KSDK-TV St. Louis, MO	10/18	Mike Wendland
Downloads	41 News This Morning KSHB-TV Kansas City, MO	10/17	Mike Wendland
Tips on sorting the growing mounds of resources on the Internet	News 2 Houston First at Five KPRC-TV Houston, TX	10/16	Mike Wendland
Surfing the Web	CNBC	10/16	Mike Wendland
Internet	KMOL News 4 KMOL -TV San Antonio, TX	10/15	Mike Wendland
Internet Options	News 5 at 5:30 WLWT-TV Cincinnati, OH	10/14	Mike Wendland
Information Overload:	Net Surfin'!	10/14	Mike Wendland

Headline:	Publication:	Date:	Author:
The World Wide Web keeps growing	pcmike.com		
The Daunting Task of Storing the Web	Internet Week	10/13	Chuck Moozakis
Cyber Historian	Communique	10/13	Alex Cramb
Wayne Business & Finance: Information overload	Westland Observer (Westland, MI)	10/12	Mike Wendland
Wayne Business & Finance: Information overload	Plymouth Observer (Plymouth, MI)	10/12	Mike Wendland
Wayne Business & Finance: Information overload	Garden City Observer (Garden City, MI)	10/12	Mike Wendland
Wayne Business & Finance: Information overload	Canton Observer (Canton, MI)	10/12	Mike Wendland
Wayne Business & Finance: Information overload	Redford Observer (Redford, MI)	10/12	Mike Wendland
Netropolitan	New Scientist	10/11	no byline
Alexa.com	David Siegel's diary (http://www.dsiegel.com)	10/9	David Siegel
A busca colective de Alexa	Universo Online--Netvox	10/7	Maria Ercilia
Internet Companion: Old sites disappear, for now	Whittier Daily News Whittier, CA	10/6	Stan Friedman
Internet Companion: Old sites disappear, for now	Pasadena Star - News Pasadena, CA	10/6	Stan Friedman
Internet Companion: Old sites disappear, for now	San Gabriel Valley Tribune West Covina, CA	10/6	Stan Friedman
Alexa: Unintelligible Agent	Washington Post	10/3	Dan Pacheco
Future Directions -- The Search	Newsweek Online	10/97	no byline

Headline:	Publication:	Date:	Author:
The Pretext Interview	Pretext Magazine	10/97	Susan Dumett
Net Surfing's Last Gasp	Wired Magazine	10/97	Joey Anuff
Unearthing the Web's New Pathfinders	ZD Internet	10/97	Melanie McMullen
Alexa Flexes its Web Index	The Net	10/97	Wayne Cunningham
Cache This	Upside	10/97	Tish Williams

Alexa Launch Coverage

Headline:	Publication:	Date:	Author:
Net Surfing's Last Gasp	Wired Magazine	10/97	Joey Anuff
Unearthing the Web's New Pathfinders	ZD Internet	10/97	Melanie McMullen
Alexa Flexes its Web Index	The Net	10/97	Wayne Cunningham
Death of '404s'	The Irish Times	9/29	Monitor
Alexa 1.0 Adds Tools and Resources	Media Central	9/24	Sean Wolfe
Alexa Searches So You Don't Have To	PCWorld Online	9/24	Lisa Moskowitz
20-something and rich! (Interview with Jerry Yang of Yahoo!)	CNNfn digital jam	9/24	Valerie Morris
Kiss That 404 Goodbye	Wired News	9/22	Chris Jones
Alexa accessorizes browsers	C/net	9/22	Alexa Lash
Tracking the Web's beaten paths	USA Today	9/18	Elizabeth Weise
End Notes	Demo Letter	9/8	Chris Shipley
Silicon start-ups on the silver screen	San Jose Mercury News	9/8	Chris Nolan
Coming Up: Specific Net Searches	Yahoo! News	9/8	Third Age News Staff
Firm develops new way to search Web	Amarillo News & Globe - Times (Amarillo, TX)	9/7	Rajiv Chandrasekaran Washington Post
Navigate Internet to track packages	Argus Leader (Sioux Falls, SD)	9/7	No byline

Headline:	Publication:	Date:	Author:
New service makes it easier to surf the Web	Kenosha News (Kenosha, WI)	9/6	Rajiv Chandrasekaran Washington Post
Internet searches improve	Concord Monitor (Concord, NH)	9/6	Rajiv Chandrasekaran Washington Post
Internet service mines mountain of data	Tri-Valley Herald (Pleasanton, CA)	9/6	Rajiv Chandrasekaran Washington Post
Seeing the Sites on a Custom Tour	Washington Post	9/4	Rajiv Chandrasekaran
New Internet Tool Takes Selective Approach (Link to Washington Post article)	Tech Talk	9/4	Ken Rutowski
Breaking the trail	ComputerLetter (Technologic Partners)	8/25	Tiernan Ray
Alexa Highlights Web Site "Trustmarks"	Internet news.com	8/25	No byline
Alexa's Theory of Relativity	PCWeek	8/18	Jim Rapoza
Cache This	Upside	10/97	Tish Williams
Download of the Week	C/Net TV - The Web	8/18	C/Net
Alexa Internet	Network News: Information for Business and Computer Applications	8/17	Alan Gatlin
Cache This	Upside.com	8/15	Tish Williams and Michael Mattis
Alexa shows the way	C/Net (Just In)	8/13	Christopher Lindquist

Headline:	Publication:	Date:	Author:
Popular opinions: Alexa tracks page use	Rochester Democrat and Chronicle (Rochester, NY)	8/10	Compiled by staff and wire reports
Service tracks web use	The Berkshire Eagle (Pittsfield, MA)	8/10	New York Times News Service
FYI: Popularity Context	Star Tribune (Twin Cities/Minnesota)	8/4	New York Times
Scent of a Web site	Inman News Features	8/4	Inman News
The Web HotList	InfoWorld	8/4	Jim Battey
BizRate Lets Consumers Rate Sites	Inter@ctive Week	8/4	Connie Guglielmo
Internet boom just beginning, conference told	Reporter (Vacaville, CA)	8/3	San Francisco Chronicle
Internet expected to grow for years	Peninsula Daily News (Port Angeles, WA)	8/3	New York Times News Service
Internet industry 'not even close to peaking'	The Midland Reporter-Telegram (Midland, Tx)	8/2	Jon Swartz
Net Surf: Alexa's New Navigation Service	Wired News	8/1	Joey Anuff
Smart Searching at Internet World (Web Addict Section)	Web Review	8/1	Wes Thomas
Insider Profile: Brewster Kahle	CMPnet Net Insider	8/97	David Sims

Headline:	Publication:	Date:	Author:
Alexa gets you where you want to go	The Seybold Report on Internet Publishing	8/97	Matt McKenzie
Site Reviews: Alexa Internet	Media Central Digest	7/29	Sean Wolfe
Alexa: Merging Web Searching and Group Experiences	Internet Outlook (http://www.webreference.com)	7/28	Richard Wiggins
Alexa Keeps Surfers on Beaten Path	Inter@ctive Week (print)	7/28	Tom Steinert-Threlkeld
Alexa Searches While You Surf	PC Week	7/28	Jim Kerstetter
Blacks Say Internet Offers Equality	The Plain Dealer (Cleveland, OH)	7/28	From staff and wire reports
Officials See No End in Sight for the Popularity of the Internet	Post-Bulletin (Rochester, MN)	7/28	Jon Swartz (SF Chronicle)
New Service Tracks Web Use	Plainview Herald	7/28	New York Times Service
Revolution Brewing at Presidio: Alexa Internet hopes to change the way you use the Net	San Francisco Examiner	7/27	Zachary Coile
Net Growth Virtually Exploding	Austin American - Statesman (Austin, TX)	7/26	Jon Swartz (SF Chronicle)
The Internet Boom Is Just Beginning: Big business is the next frontier, attendees at Chicago trade show say	San Francisco Chronicle	7/25	Jon Swartz
Getting Even More Personal...	Interactive Services Report	7/25	No byline

Headline:	Publication:	Date:	Author:
E-commerce about to launch?	The Middletown Press (Middletown, CT)	7/25	Jon Swartz (SF Chronicle)
Outside The Box	San Jose Mercury News	7/24	Chris Nolan
As Go Surfers, So Goes Alexa	Wired News	7/24	Chris Oakes
Internet Industry Giants Snub Expo	The Desert Sun (Palm Springs, CA)	7/24	Jon Swartz (SF Chronicle)
2 Internet Giants Won't Show at Expo: Microsoft, Netscape are skipping Chicago	San Francisco Chronicle	7/23	Jon Swartz
New Ad-Backed Service to Monitor Context of Web Pages	Internet Advertising Report	7/23	Compiled by Beth Cox
Service Turns Page in Aiding Web Users	The Oakland Press (Pontiac, MI)	7/23	New York Times News Service
New Guide Offers Net Directions	Inter@ctive Week (online)	7/22	Tom Steinert-Threlkeld
New Search Tool Tracks User's Path Through Internet	The San Diego Union-Tribune	7/22	New York Times News Service
News.Scan: Today's Headlines from the Mainstream Press	PC Magazine Online	7/22	NewsWatch
Looking for a Second Commercial Hit, Brewster Kahle Launches a Navigation Service	Web Week	7/21	Margaret McKegney
Alexa Internet Lands at	San Francisco	7/21	Steve Ginsberg

Headline:	Publication:	Date:	Author:
the Presidio	Business Times		
New Service Tracks Web Use	Nando.net (Reprint of NY Times Article)	7/21	John Markoff
New Service Tracks Web Use	New York Times	7/21	John Markoff
Alexa Internet: The Search as a Communal Effort	New York Times (CyberTimes Extra)	7/19	Laurie J. Flynn
Recent & Decent Web Tools	Family PC News Online	7/18	Michelle Megna
Alexa Browser Companion Makes Searches More Certain	PC World Online	7/18	Lisa Moskowitz
Re-engineering Surfing	San Jose Mercury News Online (link to TechWeb story)	7/18	Patricia Sullivan, Online Editor
Internet Historian Creates Tool to Inform Web Site Visitors	ZDNN-- The ZDNet News Channel	7/18	Renee Deger
Alexa Debuts Search Engine That Learns from Its Users	ZD Internet MegaSite Magazine	7/18	no byline
Startup Seeks To Synthesize Web	TechWeb	7/17	John Gartner
Alexa Internet Introduces Web Navigation that Learns from People	LA Times, BizWire Section	7/17	Press Release
Alexa Makes the Web History	C/Net	7/9	Alex Lash
Personal Access: Information abounds on the Internet, which means virtually anyone can find out virtually anything about someone else	Tucson Citizen	6/30	Leslie Miller

[Home Page]

Once Sojourner landed on Mars, one of the important directives that its Earth-bound project engineers had to relay to the rover each day was which way to head out. Specifically, they had to determine how it

could most efficiently reach the coordinates of its destination. To that end, engineers created a trajectory for the rover by evaluating a complex matrix of images captured on a terrain panorama, coupled with data compiled from multiple hazard-detection devices.

Navigation—to NASA at least—is truly a science and not something that can be left to chance or circumstance.

Building on that theory, sojourners on the Web could also benefit from keen navigation devices. These come in two forms: intuitive client-based search tools, and on the server side, clear pointers within a site that illuminate where users should go upon arrival. While you can't always plow a path for users through your Website, at the very least you should open all the hatches into the site and funnel traffic into the areas you want explored.

Fortunately, great strides in technology have unearthed a bevy of pathfinders that address both the client side and the server side of the navigation equation. On the client side, one smart new product is Alexa Internet's navigation service (www.alexa.com). Functioning as a tiny helper app for your browser, Alexa provides a stream of relevant recommendations of where to go next. As the service matures, so will the quality of the recommendations, as they are derived from Alexa's ongoing analysis of traffic patterns—proxy caches and anonymous usage trails. In essence, Alexa becomes the Nielsen system of the Net, as it harnesses the collective wisdom and taste of the community.

As clients move smartly from site to

NAVIGATION IS TRULY A SCIENCE AND NOT SOMETHING THAT CAN BE LEFT TO CHANCE.

site, those on the server side also need to know how users are rummaging through Websites. Traffic analyzers, entering their savvy second-generation, are a must-have investment. These tools now go far beyond the static logs of the early days. Instead, they generate dynamic tables and graphs compiled from sophisticated path analysis and user interaction.

For example, Interlogue Communications (www.interlogue.com) Whirl evaluates the movement of users across a site and identifies the optimal path to a destination. By examining users' entry and exit points, products such as this can help you understand what brought people to your site and what caused them to leave.

Accrue Software Insight (www.accrue.com) offers similar live navigational acumen. During a recent demonstration, Insight allowed us to watch a user plod through the Harley-Davidson site for four hours, all during the day while he was logged on from his government research lab address (don't tell his boss). Insight re-

Unearthing the Web's New Pathfinders

vealed where he had been before his motorcycle excursion, how long he waited to view bike pics, and how successful he was in downloading each page before Web attention deficit disorder kicked in.

These are only two examples of the many smart products available, and there are hundreds. Use them in tandem with another "live" tool—the experiences of others who have managed to capture virtual customers. In "Secrets of Successful Web Stores" on page 91, Jaclyn Easton describes how Ken Crane leveraged his Southern CA-based laser disc chain to build an e-business that has grown 600 percent since 1995. Crane's distributor, Martin Greenwald, sums up the secret of Crane's success—he knows his customers and their Web-buying ways. And that's an arduous task with anonymous, virtual shoppers. Adds Greenwald: "He's like Wal-Mart. He makes money through traffic, not units."



Julianne Richelson

THE IRISH TIMES

ON THE WEB

FRONT HOME FINANCE FOREIGN FEATURES SPORT OPINION LETTERS

COMPUTIMES

Monday, September 29, 1997

MONITOR

Death of '404s'

Remember the Internet Archive (Computimes, January 27th)? The people behind the non-profit project to make a "snapshot" of the Net have just launched a very nifty Internet navigation service. Called Alexa, the free software (at <http://www.alexa.com>) gives a toolbar for finding related Web sites, and other icons tell you information about a site - who it's registered to, how many pages it contains etc. The tool also includes an "instant messaging service" to see if your friends/colleagues are online.

Best of all, though, Alexa ties in to the Internet Archive itself. So if you get one of those "404 - document not found" messages, the Alexa server delivers an archived version of the outdated page (if it has it). Currently only available for Windows 95 and NT users, but Windows 3.1 and Mac versions are on the way...

NUA DEAL: The American Export Group has given Irish Internet consultancy Nua the contract to develop a Web site for a database of 47,000 US export companies in eight languages. "To win such a contract in the face of heavy competition from American Internet developers proves that Nua is becoming a serious player in Internet development internationally," says Nua's Gerry McGovern.

ONLINE SHOPPING: Ireland On-Line is to sell its Internet starter kit in Dunnes Stores outlets nationwide. The £4.99 pack includes a free month's full access to email, the Web etc and connection software. It will be sold in over 500 retail outlets including Dunnes, Xtra-Vision and HMV.

APPLE HARVEST: Apple launched its new OS 8 Macintosh operating system in Ireland on Friday, and will

National
Deposit
Brokers

National Deposit

Brokers

Interest rate &
Investment info

**Hamilton
Osborne King**

Hamilton Osborne

King

Irish Properties for
Sale



Premier Group

Hotels Apartments
Offices

be displaying it at the Apple Expo in the RDS next Wednesday and Thursday. The company has already sold over 1.5 million copies in the US, and it has a user base of 60,000 people in Ireland.

GAME OVER: IBM's Deep Blue, the supercomputer that beat world champion Garry Kasparov in May, has gone into retirement. A less powerful version called Deep Blue Junior will still play demonstration games, but Kasparov said from his home in Moscow that he was "very disappointed" by the announcement. He had challenged IBM to a rematch.

EMAIL PROTECTION: The head of the European Commission's "free flow of information" unit, Ulf Bruehann, has called for an independent EU data protection agency. He told the International Conference of Privacy Data Protection Commissioners in Brussels that the agency would protect citizens against the increased risks to privacy posed by the growing international traffic in personal data. He said a key area to be monitored was the treatment of email.

CNTRL+ALT+DEL: Last week's Computimes said that PostGEM/IOL was the first Irish Internet service provider to join LINX. In fact Global One says most of its Irish customers, including Medianet, have been connected to LINX since 1995.

FASTER BROWSING: Intel has introduced new technology that speeds the delivery of Web pages. Users do not have to add software or hardware to their PCs because its done at the server end. Several Internet providers plan to conduct trials of its Intel Quick Web Technology next month. It uses several techniques to accelerate downloading of graphics. The software analyses Web pages, searching for images and compressing them by removing unnecessary data bits.

SPAMMERS EJECTED: Anti-spammers on the Internet say a major battle has been won after Sanford Wallace's Cyber Promotions and two other unsolicited bulk email operations were kicked off of AGIS, a major ISP in the States. Spam-hunters believe AGIS is reconsidering its open policy toward hosting unsolicited bulk email vendors.

H-P TELECOMMUTERS: Up to 40 per cent of HewlettPackard Ireland's sales staff are to work from home "or other non-office locations". The teleworkers will have remote access to HP's huge intranet which has over 100,000 users.

IN BRIEF. . . The merger of accountancy firms **Price Waterhouse** and **Coopers & Lybrand** will also create the second largest technology and business consulting firm, with **Andersen Consulting** maintaining the number one position. . . **IBM** has developed a way to substitute copper for aluminum in making semiconductors, using a patented "fusion barrier" that keeps the copper from "poisoning" the silicon. . . **Progressive Networks**, which makes streaming media tools for the Web, has bought a major Web site for film reviews, **Film.com**. The site generates two million page views a month. . . Tests by InfoWorld magazine have found that 100 per cent **Pure Java** applications run on over a dozen platforms including Sun's Solaris, OS/2, Windows NT, Windows 95, Windows 3.1 and Macs. . .

Canadian software firm Corel has had another disastrous quarter with a \$31.4 million loss on sales of \$55.8 million. A further loss of \$15-\$20 million is also expected. . . **Gateway 2000** has introduced its Solo 2300 Multimedia Notebook, priced at £2,730. . .

Super Value-Centra has awarded **IBM** a £1.5-million contract to supply a computer system to run Ireland's first centralised distribution network for chilled, fresh and frozen foods. . . Seagate has opened its new plant in Limavady, bringing its Northern Ireland workforce to 1,500 people. . .

M I C R O F I L E

In a survey by International Data Group:

- 52% of companies in Asian Pacific countries excluding Japan believe the Internet is "important" or "very important";
- 30% of Japanese companies and 30% of US companies are in this category;
- only 16% of West European companies share this view

An America Online survey found that:

- 37% of its subscribers watch less TV than they used to;
- 22% watch less video;
- 7% watch more TV and 6% watch more video;
- the Net "has had no effect on radio, no effect on magazines, and little effect on newspapers"

According to Boardwatch magazine, approximate no. of



Sep. 24, 1997
Vol. 5 - No. 187

Alexa 1.0 Adds Tools and Resources

Losing the beta blues, Web navigation service Alexa (<http://www.alexa.com>) has relaunched its service and beefed up the content available to its users through an alliance with Encyclopaedia Britannica (<http://www.eb.com>).

The new alliance aims to put reference materials at users' fingertips so they can retrieve information without leaving the Web page they're on.

After being in beta testing for roughly three months, Alexa's new version (1.0) features a fully debugged toolbar that attaches to a user's browser. The toolbar provides information about each site requested, its owners and relative traffic. If the site sought is off-line but otherwise publicly available, users can usually find an archived copy of the page on Alexa's database.

Users can access content from new partner Encyclopaedia Britannica through a pop-up browser window launched via the toolbar. The alliance pairs Encyclopaedia Britannica's 32 volumes and some 5,000 with an online dictionary and a thesaurus from Merriam-Webster to create a kind of reference shelf.

*Cowles/Simba Media Daily 9/24/97
Copyright 1997 Cowles Business Media. All rights reserved.*

[Back to Media Central](#)

Media Daily can be reached by [e-mail](#), by fax at 212-683-3986 or by telephone at 212-683-3540 x416

Copyright 1996 Cowles New Media

[Current Media Daily](#) | [Media Daily Archive](#)

Internet Archives preserves content

ARCHIVE FROM BE

The collection process is not comprehensive, he said. The text, graphics, and audio clips are limited to only what his crawlers troll, missing hundreds of thousands of sites restricted to the software spider.

This archive has already created, in affiliation with the Smithsonian Institution, the '96 U.S. Presidential Election Web Archive. Here are kept the original political home pages, campaign Web sites, and costly critical election '96 coverage links that forever mark the political collision of politics and cyberspace.

"Unless someone starts saving it with every day that passes we're losing the record of one of the great turning points in human history."

Kahle's idea to build a freely accessible, nonprofit archive didn't stop with the huge computers that store the terabytes of data. The Internet Archives' brother is a commercial arm called Alexa Internet.

Alexa is The Internet Archives' private for-profit attempt to aggressively develop software for manipulating terabyte-level data. It is part of his commercial aspirations, he said, to sell the Web searching and storing technology developed at the archives.

The service — called Alexa after the ancient library at Alexandria, Egypt — doesn't rely on a keyword search premise.

Instead, Alexa, a software add-on for Web browsers that can be downloaded from www.alexa.com, uses several powerful computers to crunch through the archived data in a process called data mining. It looks for patterns within the Web, ultimately providing a few recommended strands to travel.

If a surfer should hit a "dead" Web site, Alexa offers to retrieve the expired page from The Internet Archive.

"We've only just been getting warmed up," said Kahle, of his \$2 million personally-bankrolled investment that he admits has given him more payback as a labor of love than a financial boon.

Kahle left Boston-based Thinking Machines in 1992 to start an Internet publishing company called WAIS Inc. It was sold to America Online Inc. of Dulles last year for about \$15 million.

He is constantly retooling Alexa to be more user friendly and someday allow surfers to scour the vast reservoirs of the archive more effectively.

However, the task has its detractors. Some view his quest as a Quixotic journey into the realm of the impossible and absurd.

"I think it's laughable," said Stuart Mantel, technology project manager for Time Inc.'s Pathfinder Web site. "What are they really going to accomplish? Is the Web really just the sum of its parts?"

"I think it's laughable. What are they really going to accomplish? Is the Web really just the sum of its parts?"

STUART MANTEL
TECHNOLOGY PROJECT MANAGER FOR
THE INC. STRAIGHT PATHFINDER SITE

Video conferencing, streaming audio, quick-ram interaction, and more are part of the truly ephemeral aspects of the 'Net that can't be stockpiled, he said.

Mantel comments Kahle's quest as noble, but asks, "If you plunge a bucket into a river, do you pull out the essence of the river — or just a gallon of water?"

Despite his opinions, tributes to Desert Storm, Mars Explorer, and QJ Simpson on Mantel's Pathfinder site have already been deleted and stored in the inaccessible computer-labeled archives of Time's Web site.

But critics are the least of the formidable challenges to The Internet Archive, starting with traditional copyright and privacy law.

When a marriage goes sour, and an archived home page that shows the couple smoothing can't be torn up like the hard-copy wedding photos, should the divorcees have a right to remove them? Should a U.S. senator be able to erase data posted from his or her college years?

And does collecting information made available to the public violate the "fair use" provisions of the copyright law? The issues are not easily resolved.

Kahle admits these issues get sticky and said he is working with major policy makers and experts on intellectual property to help understand the scope of the copyright issues the company will soon face.

For now, to address these worries, Kahle lets authors exclude their works from the archive if they choose.

Middlesex News
Framingham, MA
September 21, 1997

Preserving a moving target

*The Internet Archives hopes
to save Web content – all of it –
for posterity*

By Tom Spring
NEWS TECHNOLOGY WRITER

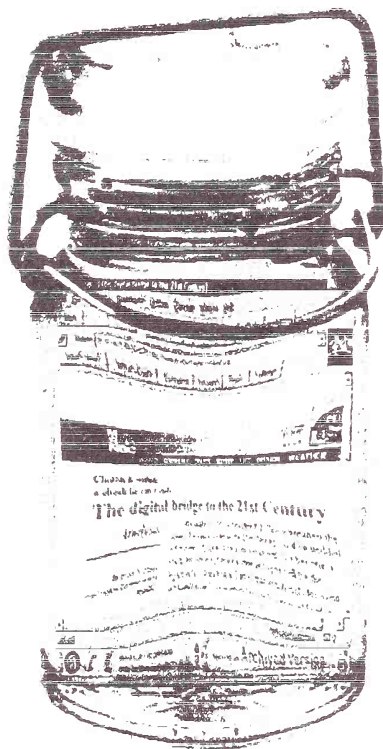


PHOTO ILLUSTRATION BY JEAN BEECHER
The Internet Archives' www.archive.org has downloaded

The World Wide Web is a perpetual cycle of here today, gone tomorrow. The Internet you log onto one day can vanish the next, lost in the cyber-busyness of digits and data.

Yet, as cyberspace grows to be a public forum where homepages, Web zones, and scholarly essays reside, the Net becomes a reflection of our digital heritage, experts believe.

And like the first

books, films, and even newspapers now decayed and unrecognizable, history threatens to repeat itself with the World Wide Web.

Much has been said about archiving the Internet, but little has been done – until recently.

About a year ago, a nonprofit organization cast the first footprints of the Information Age to be preserved forever. Thanks to computer scientist and millionaire entrepreneur Brewster Kahle, the Internet – from the sublime to the ridiculous to the brilliant – will be preserved in its chaotic exactitude.

Known as The Internet Archives, the organization has quietly begun the daunting task of creating the first and only known archive of the Web.

"It's important that we are able to look back," said Kahle, who invented the pioneering Net

are out of print. And the same holds true for the Internet.

The average document on the Web has lifespan of 75 days, and for this reason historians and technology professionals, like Microsoft's Chief Technology Officer Nathan Myhrvold, believe Kahle's project is vital in a medium that is fleeting.

The goal is to have an unprecedented permanent record of all publicly accessible data, whether it's a neighborhood home page or a foreign-language Usenet posting about Japanese haiku, he said.

This cultural time capsule of the digital revolution is housed in a space no bigger than an average living room. So far Kahle has stored four terabytes of the Internet on a combination of tape and hard drives in his San Francisco offices overlooking the Golden Gate Bridge. To give you a sense of breadth of the library, four terabytes is equal to 4 million megabytes. And despite the fact that the Net doubles in size every six months, according to Kahle his digital archive is ready to stuff the whole thing into his offices for decades to come.

Not that it's all in demand.

*"Unless someone
starts saving it with
every day that passes
we're losing the record
of one of the great
turning points in
human history."*

BREWSTER KAHLE

"Most of it is junk," he admits. He views his records not as a data morgue, but as a living history. Between CNN.com, the alt-textless-jokes newsgroup, and the home page for The Dallas Cowboy Cheerleaders, Kahle believes he has captured the zeitgeist of the Web circa 1990, for historians to ponder centuries from today.

The logistics of taking snapshots of the Web are relatively simple. The software on Kahle's computers "crawls" the Net, downloading millions of Web pages from one site after another.

Once a page is captured, the software looks for cross references or links to other pages. The software then makes copies again and sends additional links contained in the new pages

Site Spotlight

Who can keep up with the dirt on The Bold and the Boob-Jobbed in Hollywood? If you're going to bite the bullet and take this stuff seriously, it's time to try these entertainment news sources.

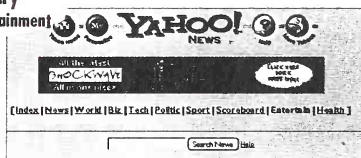
The Hollywood Reporter
www.hollywoodreporter.com



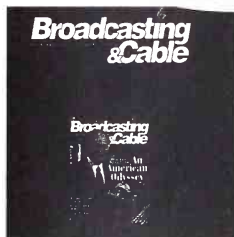
The Reporter claims the dubious honor of being the first in a venerable line of Tinseltown trade rags. You can peek at showbiz's top headlines, but you have to subscribe to get to full stories, the archives or film and TV production listings.

Yahoo Entertainment Summary
www.yahoo.com/headlines/entertainment

A metapage for entertainment information, Yahoo's daily entertainment news site draws stories from *People* magazine, *Variety*, Reuters and dozens of other sources.



Broadcasting & Cable Online
www.broadcastingcable.com



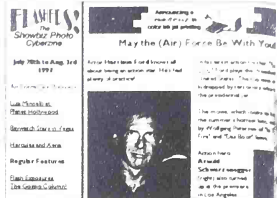
Find out just where things stand with digital TV, cable modems, telephony and rate regulation at Broadcasting & Cable newsweekly's teaser site. Our favorite quote, from that semantic genius Al Gore, who is pushing to make the ratings symbols larger on TV's sauciest shows: "Today, America's parents have won back their living rooms."

TV Guide Entertainment Network
www.tvguide.com

Who says the Web is just a vehicle for the tony digital elite? Rupert Murdoch's online *TV Guide* is just as aggressively mid-brow as the print version, and includes the same tired gossip from ubercolumnist Liz Smith and "news" on your favorite idiot-box cheesecake and beefcake.



Flashers
www.flashers.com



Dig up the dishiest dirt on your favorite supercheesy TV shows and new movie premieres, from "Baywatch" and "Xena" to "Men in Black." Copious photos of today's Hollywood "talent," from Sherry Lansing to the babes from "Baywatch."

Cache This

www.turtles.org/happen.htm#logger

This site tracks a turtle's migration from California to Japan in a way that opens science to students—an example of sensors getting integrated with computer networks in interesting ways. If this kind of thing catches on, then maybe life doesn't have to be a spectator sport.

www.epic.org

The Electronic Privacy Information Center's aggressive but sane presentation of one of the trickiest and most important social issues happening in the technology sphere.

pharmdec.wustl.edu/juju/surr/surrealism.html

The power of Web publishing—inexpensive, highly interconnected and equalizing—allows for sites like this one on surrealism.



Brewster Kahle
CEO, Alexa Internet
www.alexa.com

◀ PREVIOUS

▶ NEXT

| CONTENTS

PRETEXT MAG

THE PRETEXT" INTERVIEW

“Brewster Kahle
talks to
Susan Dumett.

We're archiving
the link
structure of the
whole Net.



During the last decade or so, Brewster Kahle has designed supercomputers, invented WAIS (Wide Area Information Servers system), and founded the Internet Archive (see [The Web Museum](#)). His newest venture is Alexa, a piece of navigation software that keeps track of the paths previous visitors followed from a given Web site and then makes recommendations for additional sites to visit based on their choices. Available for free from the [Alexa Web site](#), it just may change the way information is organized and distributed on the Internet.

From his office near San Francisco's Golden Gate Park, Kahle spoke by phone with PreText writer Susan Dumett about Alexa and the future of digital libraries.

Susan Dumett: How does Alexa work?

Brewster Kahle: Alexa is based on the collective wisdom of users who've surfed the Internet before you. Alexa asks, "Where did other people go after visiting this site? Let them find the quality and weed out the hype." It's different because it's a companion—it's with you on your screen as a tool bar providing information that changes based on where you are and where you might want to go next.

So, if I'm at an astronomy site, Alexa could refer me to sites other people have visited from this site that may or may not be related to astronomy?

As the number
of sites grows,
I'm not sure
we can expect
directories to
keep up.

There is quality
out there. It's
buried among
tons of
irrelevant stuff,
but it's out
there.

be related to astronomy.

Well, you're not going to get a sports site because you're not going to get many people reinforcing that particular subject. We're archiving the link structure of the whole Net. We can find sites related to the page you're looking at--not based on the links on the page, but by what other people have thought of that page. It's kind of a "have you read any good books lately?" form of navigation.

Wired magazine recently ran a short piece questioning Alexa's value as people lose interest in surfing the Net. But Alexa seems not so much a surfing tool as an organizational tool--performing the role of cyber-librarian in a sense.

Well, librarian isn't the most exciting metaphor, but it is filling that role. Often the value of being in a library is not the book you're looking for, it's the other books on the shelves. We're organizing those shelves. Alexa's not a surf medium, but a medium for finding things.

When you're over looking at that particular book on the shelf, we show you the other ones around it--not just any old books or even similar books, but ones that other users have thought were important. The number of Web pages is now over 100 million. Finding the right information requires pretty sophisticated techniques. We think the technique that's going to work is leveraging the intelligence of the users. It's not like we have the smartest robots here. We think people are the smart part. We help people leverage what millions of others have done.

What do you mean?

The number of Web sites doubles every six months. What are the current technologies for finding things right now? There's directories that have editors type in what they think are the best



What happens to an idea that
wins a NOBEL PRIZE?



washingtonpost.com | [home page](#) | [site index](#) | [search](#) | [help](#) |

Technology Post: Your plain-English guide to the digital world.

Web Exploration: learn your way around the Web through guided tours.

Technology News and Analysis: explanations of issues in the digital realm, plus updates from The Post and the Associated Press.

Alexa: Unintelligible Agent

By Dan Pacheco

Friday, October 3, 1997; Page N70
The Washington Post

WE web users are a fickle breed. On the one hand, we want easy access to any information, no matter how obscure or useless. On the other, we resent having to dig around to find it. In an ideal world, the Internet would figure out what we're looking for before we think of it.

Hence the rise of supposedly "intelligent" programs that attempt to anticipate your next click. The latest is Alexa, a Web copilot that follows you around and recommends sites it thinks you'd like.

Alexa makes its decisions in two ways. First, it looks for similarities among the site on your screen and others stored in the Internet Archive -- a compendium of (theoretically) every Web page in existence, compiled by its developers. If, for instance, this analysis indicates most Web sites about deer also mention rifles, Alexa may recommend a hunting site to someone reading about whitetail migration patterns.

The second method is called "collaborative filtering." As people travel the Web with Alexa, their tracks are recorded anonymously in a master archive. Users can also vote for sites they like and trash ones they hate, further padding the preference list. But since not many people have used Alexa yet, this feature isn't all there.

After a week-long trial, that was obvious. Like an annoying store clerk, Alexa was continually pushing the equivalent of ugly slacks and sweaters when all I really wanted was a pair of black socks.

At its best, Alexa bumped me from the University of Colorado's Web site to schools in Hawaii, Kansas and Illinois. At its worst, it made bizarre connections between disparate subjects; I still can't figure out what Miller

Genuine Draft has to do with an Internet service provider called "The Human Factor."

It only took me a day to realize that Alexa's limited universe of suggestions and categories paled in comparison to any of the hand-picked, more populated categories of Yahoo or the software-generated, but still useful LiveTopics feature of Alta Vista .

More interesting are the details that Alexa can display about a site. If you're feeling nosy, click "Where am I?" to pop up a tiny window with a one-to-five-star rating, the name and address of the site owner, the relative speed of the site over time, and the gross revenue of the owning company in 1996. This can be enlightening; I had no idea that washingtonpost.com, the Web site I've worked on over the past year and a half, contained 38,000 pages.

Elsewhere in Alexa, a book icon brings up a search form for Encyclopedia Britannica Online and the Miriam-Webster Dictionary . A telephone icon supposedly lets you talk with other Alexa users, but in a week I never came across one chatter.

But it's the smallest icon, curiously tucked away at the right-hand corner of the screen, that holds the most promise -- and delivers the most disappointment. If you ever come across a missing page ("Error 404: File Not Found"), clicking this icon will retrieve a copy from the Internet Archive itself. Unfortunately, Alexa only delivers one randomly-chosen version, usually around six months old.

More useful would be a feature that lets you retrieve every version back to a site's beginnings -- something the Archive should have no problem doing. Then again, who said there's any such thing as a free lunch?

Alexa, Win 95 (requires Netscape Navigator 2.0 and up or Microsoft Internet Explorer 3.0 and up), free at <http://www.alexa.com>

© Copyright 1997 The Washington Post Company

[Back to the top](#)

washingtonpost.com

| [home page](#) | [site index](#) | [search](#) | [help](#) |



What happens to an idea
that wins a NOBEL PRIZE?


 Alexa

Net Surfin with Mike Wendland

Welcome to Net Surfin with Mike Wendland. Here you'll find all the sites that Mike features on his weekly segment seen on CNBC's "Steals and Deals" and sent out to all 215 NBC stations in the U.S. by the NBC Newschannel. An archive of previous Net Surfin' stories also is available.



Check out Mike's high-tech special reports. Latest story: **Beware of Virus Myths**.



Notice! Mike's High Tech pages now use Flash! animations for its top navigation bar. If you don't see the navigator, make sure to download the plug-in today!

The Net Surfin' Theme is an original composition by Dan Bowyer.

Join Us!

Find answers to your computer questions, talk about high-tech issues and join the High Tech Talk/Net Surfin online community! It's all at Mike's Bulletin Board. Try it today!

Latest Selection: Alexa

It's called Alexa ... named after the fabled Library of Alexandria in ancient Egypt ... the world's first and last attempt to catalogue it's collected knowledge. And in that same spirit this free, downloadable program tries to do the same thing on the World Wide Web.



After you download it, it lies at the bottom of your screen whenever you're on the Web... The Alexa toolbar is ready to tell you who owns the site you're visiting, other similar sites you may want to check out, and how popular it is, or how many hits it receives.



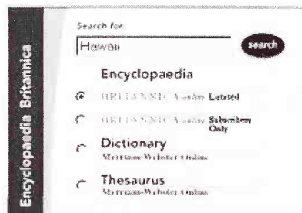
There's also an instant messaging system that allows users to communicate with other Alexa people, too.

But my favorite feature is an instantly accessible link to Encyclopedia Britannica reference tools ... like the dictionary and thesaurus. You can also get to the online edition of the encyclopedia, but ... *warning* ... they charge for total access, though Alexa users can get a free trial subscription.

This Alexa program is not a search engine. It's called a navigation service and what it does is add context to all that content out there, a big help as the Web continues to double in size every six months.

Do you have a favorite site on the World Wide Web that you think would interest our nationwide audience? If so, just tell Mike. If Mike uses your suggested site on-the-air, he'll send you a supercool NET SURFIN' T-shirt. Don't forget to include your name and address!

Got a comment about computers or life on the information superhighway? You can contact me anytime, online at mike@pcmike.com.



Virus Warnings - Watch Out for Hoaxes!

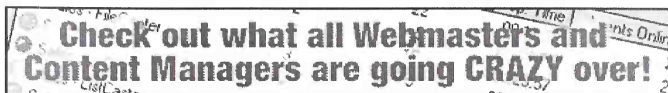
Heard the latest warning about the destructive new computer virus? Well, maybe you should think again. Before you panic at the thought of a computer virus wiping out your machine ... the next time you receive one of those e-mail "the sky-is-falling!" warnings about some supposedly "new" virus ... rush, don't walk, to <http://www.kumite.com/myths/>. Most virus scares are baseless, hysterical and unfounded.

Don't ever pass on a virus warning unless you KNOW it's real.

Don't take the e-mail sender's word for it, check it out yourself and remember, it's in the interest of the anti-virus software makers to get as many people worried as possible.

Copyright © 1997 WDIV. All rights reserved.

[PCMike.com](#) | [Special Reports](#) | [High Tech Talk](#) | [Previous High Tech Talk Stories](#) | [Tech Tips](#)
[Net Surfin](#) | [Previous Net Surfin Stories](#) | [Hot Links](#) | [Mike's WWW Board](#) | [WDIV Online](#)



PCWORLD Software

Join Now...	Index	Site Search	Help
Home	Software Library		

Alexa

This Web navigation service works with Microsoft Internet Explorer, Netscape Navigator, and Netscape Communicator browsers to suggest Web sites you might want to visit next based on the site you're currently checking out.

DOWNLOAD	
alexasetup.exe	1.81MB 31m *

*estimated download time
(based on 14.4 modems)

FILE FACTS	
Author:	Alexa Internet
Version:	1.0
Operating System:	Windows 95 & Windows NT
Type:	Freeware
Upload Date:	24-SEP-97
Downloads:	1901

ADVERTISEMENT

(Sorry)

CATEGORIES

[Anti-Virus](#)
[Business & Office](#)
[Desktop Utilities](#)
[Entertainment](#)
[Essential Shareware](#)
[Games](#)
[Graphics](#)
[Help Files & FAQs](#)
[Internet & Comm](#)
[Sound & Audio](#)
[Programming](#)
[System Utilities](#)
[Video](#)

FREE NEWSLETTER

Sign up for your FREE
PC World Shareware
Pick of the Day!
 email address here

SUBSCRIBE

QUICK SEARCH

Power Search

Many more search
options!

TODAY'S HEADLINES

[IE4 Security Flaw
Fixed](#)

[Netscape's Netcenter
to Include Electronic
Commerce Services](#)

[More Daily News](#)



Return-Path: <staff-request@alexa.com>
X-Sender: brewster@mail.alexa.com
Date: Thu, 06 Nov 1997 10:42:16 -0800
To: cool-stuff@alexa.com
From: Brewster Kahle <brewster@alexa.com>
Subject: Alexa as a part of the next wave!
Cc: cynthia@alexa.com

By Mitch Ratcliffe
September 18, 1997 4:52 PM PDT
ZDNN

After the search engines come the syndicators, companies looking for value in reusing, promoting and linking sites and content to drive value. Keep an eye on this new breed of Web content aggregators.

Several interesting examples of the different styles of syndication are arriving on the Net this month, including NetEvents Rapid Content, Alexa Internet's archive and "where to go next" features, PointCast's and Netscape's/Individual's targeted vertical information channels, and the evolving Mining Company.

Even the Electric Library and Northern Light, two search engines that incorporate extensive archives of articles from the paper press, fall into this category, since they try to extract more revenue from reused content.

I'll go out on a limb here: This category represents the next source of explosive growth on the Net, and the path fraught with the greatest peril for startups.

Remember, the real money in television is made in syndication. Sheldon Leonard, William Shatner and Bob Denver made their fortunes on the thousands of reruns of their programs. First-run shows seldom earn back their production costs. Can the same system work on the Net? The answer is a big "yes."

Syndication is the art of picking the right content for an audience that can be reached at the lowest possible cost -- and that cost includes creative and marketing costs, not just the raw expense of network connectivity.

On the Net, that means the syndicator has to begin to reuse content almost at the moment it is posted.

Obviously, this presents some problems, what with copyright and other nasty intellectual property issues still so up in the air. But, because the Net allows content to be reused to fill other sites without actually creating copies of the original material, there is ample opportunity to capitalize on the floods of information merely by directing it into manageable channels.

The question is whether the syndicator has the right model. Three basic strategies are emerging, which I'll examine in my upcoming columns. In a nutshell, the syndicator business models are:

- 1) Branded sites that deliver a collection of materials from other sites. These sites add value through editorial or personalization features -- either they sort through the Web for content appropriate for their audience or they build sophisticated back-end systems that filter information for the user. This group includes PointCast Editions, Netscape/Individual channels, and most news sites, including CNN and our own dear ZDNet, which use third-party sources along with their own content.
- 2) Pass-through sites and servers that deliver aggregated content streams to a third-party Web site that puts its own brand on the repackaged content. This is the NetEvent strategy and The Mining Company's corporate channels strategy, the first example of which was with CitiBank.
- 3) Omnigators, sites that collect and personalize information from across the Web and can deliver services through their own sites or client

software, or through third-party sites. These sites include Alexa and the consumer side of The Mining Company, which uses human guides to collect and deliver content from across the Net.

Each of these strategies could be successful, if the provider is careful with its resources, whether human or computational. The problem for all is growth and scalability. As we've seen throughout the history of the Web, exponential growth can destroy the best laid plans.

THE NEW SYNDICATORS, PART II

Having laid out the foundations of the Net syndication market, let's take a look at the challengers to the crown. As I said in my last column, the real trick to syndication is combining a low cost with a high volume of exposures in order to drive revenue.

I've grouped the sites according to the general strategies they follow. Some are branding themselves, while others are planning to charge third-parties to place a brand on their content. Finally, one is taking the big leap, trying to sweep everything on the Web.

You'll notice that The Mining Company shows up under two categories. It's the most comprehensive approach to the Web syndication business, and one of the strongest; consequently, it also faces the greatest hazard of syndication -- can it keep up with the demand for more human and business relationships to keep its business afloat?

Branded sites

- * PointCast Editions. The paragon of a broadcast channel on the Web, PointCast will rely on third-parties to deliver aggregated content that deals with specific industrial categories. EDS, a former General Motors subsidiary, for instance, will handle the automotive industry news. As long as PointCast can find partners, its costs will be manageable. And if they can't find partners to take on a new "edition," they are not compelled to grow.

At 375 million hits a month, the basic service is a formidable foundation for growth through segmentation into attractive markets, which is the very soul of Web syndication.

- * Netscape Netcaster vertical reports. Like PointCast, Netscape and Individual Inc. hope to attract a loyal audience from the alleged 60 million Netscape Navigator and Communicator users on the Web. And, also like PointCast, the venture's ability to grow is limited only by the capacity of its partners to bring content and context together.

But, by turning the Netscape home page into a fee-based grazing pasture for big content providers, Netscape runs the risk that it will alienate the smaller content publisher, who they also need as server customers. The answer: a PointCast Connections-like channel that gives anyone access to the Netscape home page traffic.

- * Electric Library. One of the best search sites on the Web, Electric Library's real business is selling library cards that grant unlimited access to an attractive archive of information from expensive journals and photo sources.

This can be a very lucrative business, if their collection continues to grow and remains an attractive alternative to the free stuff on the Web. At \$60 a year for a subscription, plus advertising revenue, this is a syndication move that has legs.

- * Northern Light. The one-off sales version of the Electric Library, Northern Light maintains its own extensive library of content that you usually can't find on the Net, like Foreign Policy and New Perspectives Quarterly.

Northern Light will sell access to individual articles, along with a better-than-average search engine that sweeps in the contents of the entire Net. With advertising revenue and a loyal following, which the best search engines have proven they can build, Northern Light is perhaps even more profitable than Electric Library can be, since a customer can spend and spend all year long.

* The Mining Company. The consumer side of The Mining Company site brands human guides who lead visitors to the best of the Web. It's a unique editorial play, but one that is very difficult to scale.

However, because The Mining Company is pursuing several lines of business simultaneously, using the same content in different packages, it has the potential to create the highest value from each piece of information syndicated by its guides.

Pass-through sites and servers

* NetEvents Rapid Content. A pure syndication play, the NetEvents team would like to be the King World Syndicate for the Web, placing its content in a wide variety of sites, just as cartoons are distributed for use in many newspapers.

NetEvents' customers are sites that want to blend information from many sites to create something that draw customers back again and again. That means the company has to find the best of the Web, establish a profitable relationship with the provider of that information, and then sell third-party sites on the desirability of the content.

Its first relationships revolve around the well-worn personal investment information category, so it's hard to tell whether NetEvents really has the vision to open new markets for information syndication.

* The Mining Company. It's deal with CitiCorp, which will populate banking sites with content drawn from its guides sites, like investing and small business consulting, is another pure syndication play. At more than \$3 million for the CitiCorp deal, this can be a very good business.

Omnigators

* Alexa Internet. An archive of the entire Net, with extensive recommendation and rating capabilities, the Alexa service is an odd hybrid of collaborative filtering, buddy list communications and Web navigation assistance.

Alexa attempts to offer users a comprehensive, yet personalized, guide to the Net, a very difficult trick to pull off. The facility of the client/server technology Alexa uses is unproven, though founder Brewster Kahle assures that the system "learns over time and gets better with use."

If it works, it could become a de facto interface for mapping the relationship between different pages on the Web. As a syndication business, the ability to provide context anywhere on the Web could be the most powerful differentiator of all.

Next week, in part three, I'll look at the winners and potential losers in the battle of the syndicators.

Mitch Ratcliffe is president of Internet/Media Strategies Inc. (www.ratcliffe.com), a Tacoma, Wash., consultancy. He can be E-mailed at godsdog@ratcliffe.com.

THE NEW SYNDICATORS, PART III

There is more than one way to skin a cat, and more than one way to syndicate the Net. Several of the companies we've been discussing in my past two columns will succeed, if they can combine quick responses to changing business situations with a scalable technology.

Human skills and human resources will be the deciding factor in the syndication race. The technology for doling out information is plenty mature, so there's not a lot of differentiation to achieve, but in the realm of editorial value the opportunity to add value is almost unlimited.

In an environment filled with information, editorial and business development savvy can be applied to building syndication channels that serve from 100 to 100 million people profitably.

I think The Mining Company has the right basic idea, though I suspect it has a lot of work to do on implementation to make itself succeed. The guide concept, that a person with interest and expertise in a particular area is the best filter for that area, is a great one.

Moreover, the company is distributing its human resource costs across the Net. It needn't hire people, it merely needs to convince potential guides that it can market and pay them better than others.

But, the two key problems at The Mining Company will come as consumers and business look for contrasting opinions -- currently, the company assure guides exclusivity in their areas of interest, and the best guides will recognize their increasing value and demand more money.

The exclusivity problem can be solved by introducing meta-guides, who assemble several guides' sites into a multiple-perspective online publication; but this costs money, up-front for editorial employees. This will be crucial to the company's corporate services effort. How they will deal with guides making demands for more money, or a guide union, is anyone's guess.

The NetEvents strategy, to syndicate content through its servers to third-party sites, is the cleanest and most scalable of the companies in this article. They needn't build a lot of sites, just shovel the content and value-added features, like reader rating systems, through to the third-party site, where it will appear under another brand. The content is still served off the original providers' sites, so NetEvents will not have to expand its server resources to support more content.

"This allows us to scale quickly by focusing on very efficient distribution techniques rather than building the mother of all centralized networks," said Joel Maske, CEO of NetEvents.

Making the most of all those connections, though, is a very tough job. NetEvents must develop very refined packaging skills, or there's no reason for content providers not to compete against them for hits. The company will put 20 percent of its resources against finding and making deals with the Web's best content providers. Most of that should go not to Web surfers ala Yahoo!, but for the first generation of Web talent agents and acquisitions editors.

Electric Library and Northern Light, search engines combined with fee-based special collections, may find themselves hamstrung by costs. In order to succeed, they have to overcome the popularity of other search engines.

As a recent NPD Group report found, search engine users, particularly Alta Vista users, are very loyal and visit the engines frequently. For either Elibrary or Northern Light to compete, they have to win users through better overall search capabilities. That means they have to build out their systems to support more traffic, and those extra people may not necessarily pay enough to support the added hardware. Elibrary definitely enjoys an advantage, since it charges an up-front membership fee.

PointCast and Netcaster, the old hands in this game, will confront both scale and business relationship problems as their syndication businesses grow. PointCast, though, is far better prepared to handle the argument that they have begun to compete against their content-providing customers on the regular PointCast network -- if someone complains loudly, PointCast can simply promote them to a PointCast Edition, one of the syndication networks.

Netscape, as I've said elsewhere, will certainly alienate content providers by pricing space on its popular home page out of many smaller companies'

reach. That may cost Netscape in the server market, as its customers turn to other "non-competitive" servers (after all, who wants to subsidize competition for readers by buying a server? Well, besides Microsoft customers?).

Finally, Alexa, the omnigator syndication engine, is really on to something with its catalog-everything-and-map-links-between-sites strategy. However, in order to ensure availability of data, it has to store everything it finds on the Net. This means it is creating copies of sites, which raises all sorts of copyright flags. Someone will sue Alexa, and soon. Alexa has technical workarounds, but those won't prevent the lawyers from striking. The company's success ultimately depends on its copyright strategy and its lawyers abilities.

On a scale of one to ten, The Mining Company and NetEvents earn a 7.5, PointCast a 7.0, Electric Library a 6.0, Northern Light a 5.0 Alexa a 4.0 and Netscape Netcaster a 2.9. So, place your bets and enjoy the syndication ride.

Mitch Ratcliffe is president of Internet/Media Strategies Inc. (www.ratcliffe.com), a Tacoma, Wash., consultancy. He can be E-mailed at godsdog@ratcliffe.com.

-----End forward message-----

Canton Observer
Canton, MI
October 12, 1997

WAYNE BUSINESS & Finance

Information overload

World Wide Web keeps growing and...

I don't know about you, but I'm getting tired. The information overload on the Internet has become simply unmanageable. I spend two hours a day online just keeping track of new sites. And every night when I log off, there are still a couple dozen more that I don't have the time to click.

Another survey came out this past week that reported the number of Web sites on the Internet is continuing to double every six months. Terrific.

But I did come across one site this past week that offers relief and convenience.

It's called Alexa (www.alexa.com) and it's named after the fabled and long lost library of Alexandria in ancient Egypt, the first and last attempt to archive and catalogue the collected wisdom of the world. In that same spirit the free, downloadable program you can pick up at the Alexa site tries to do the same thing on the World Wide Web.

MIKE
WENDLAND



PC
TALK

For more than a year, the people at Alexa have been copying every Web site they come across. As of last month, they had more than seven terabytes of information on more than 500,000 separate Web sites. A terabyte is a million megabytes.

Know how big that is? Consider that all of the books in the average branch of a public library contain about three terabytes of data.

That massive Internet archive they've collected is used by the Alexa program to provide relevant recommendations on where to go next on the Web. They call it "intelligent navigation."

The download installs with

just a mouseclick and thereafter, it lies at the bottom of your screen whenever you're on the Web.

There's also an instant messaging system that allows users to communicate with other Alexa people, similar to America Online's Buddy notification system or ICQ from Mirabilis (www.mirabilis.com).

My favorite Alexa feature is an instantly accessible link to Encyclopedia Britannia (www.eb.com) reference tools, like the dictionary and thesaurus. You can also get to the online edition of the encyclopedia but... warning... they charge for total access, though Alexa users are offered a free trial subscription.

Alexa only works with Windows 95, Macintosh and Win 3.1 versions won't be available until mid-1998.

CAMP

THE NETWORKING NEWSPAPER October 13, 1997

Formerly
CommunicationsWeek

INTERNETWEEK

www.internetwk.com

Case Study

THE DAUNTING TASK OF STORING THE WEB

BY CHUCK MOOZAKIS

TRACKING THE CONTENTS of the World Wide Web is one thing. Storing the contents is quite another.

But storing the World Wide Web is exactly what Alexa Internet does each and every day.

That's part of the purpose of Alexa, a free Web navigation service launched late last month. Alexa also tracks the sites Internet users visit on the Web and, taking a page from *Consumer Reports*, grades the value of a site based on a series of customized benchmarks and other metrics that vary de-

STORING, PAGE 35 >



CEO Brewster Kahle: Alexa is like the Web's Library of Congress.

STORING

► CONTINUED FROM PAGE 31

pending upon the type of site.

But Alexa goes one step further—it archives the public component of the Web. The company also generates statistics garnered from monitoring these sites, logging such things as how frequently the sites were updated and how often the linked sites were accessed.

Alexa (www.alexa.com) can also eliminate those pesky "404 File Not Found" errors. Since Alexa has a collection of what the Web looks like daily, users can request pages even after they are no longer displayed.

In other words, Alexa is the Web's version of the Library of Congress. Whatever has appeared on the Web, Alexa can find it and deliver it automatically.

The end result, said Alexa Inter-

net president and CEO Brewster Kahle, is that people cannot only determine how valuable a particular site might be, but gauge information that has appeared on that site.

This has significance to people who may want to find out what happened on a particular day and how an Internet news service may have covered it. MSNBC (www.msnbc.com), for example, a popular news service among Web surfers, doesn't archive its materials, which means Alexa appears to be the only source to determine what news it covered on a specific day.

And MSNBC is not the only site that doesn't store its old pages, Kahle said. "As the Web becomes more of an important platform, people are going to want a historical archive of what has happened."

Alexa launched its service late last month, after six months of determining the best way to manage the massive amounts of data creat-

ed by people using the Internet. Not surprisingly, storage was a prime consideration.

"We needed to do a tremendous amount of data mining and we had to determine how much it would cost per gigabyte to store and retrieve the data," Kahle said.

Kahle and Mike Burner, vice president of development, con-

"As the Web becomes more of an important platform, people are going to want a historical archive of what has happened."
Brewster Kahle

cluded that a robotic storage library—using Digital Linear Tape media—was what they needed.

"We found it cost \$240 per gigabyte to store data on a hard drive; a tape robot was \$23 a gigabyte and the tape on the shelf is \$2," said Kahle. "We decided to go the tape robot route."

Since Alexa was tracking more

than 6.5 terabytes of data each day—encompassing 600,000 Web sites and more than 135 million pages—gigabytes added up quickly without a significant delay, according to Bob Lindsay, product marketing manager at StorageTek.

Alexa began storing data on the 9710 shortly after the \$100,000-plus system was installed in April, Burner said. A two-month beta test ironed out last-minute wrinkles, and Alexa began operations without a hitch, he said.

Alexa stores its data on 422 of the 9710's 588-cartridge capacity, using 35-gigabyte DLT 7000 units.

"We have more than 6 terabytes of data now," said Kahle. "We already have 100 times more data than AltaVista has, and the Web continues to double in size every six months. We will do what we can to keep up with the growth of the Internet, and do what we can to help people make sense of it."

terabytes of storage capacity. Aver-

Inter@ctive Week
October 27, 1997

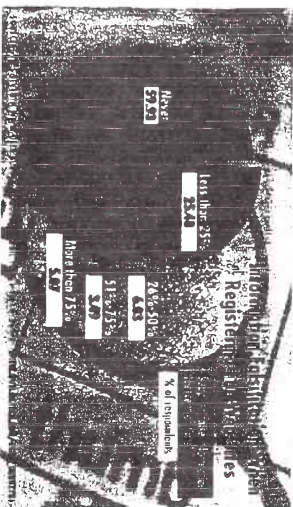
HYPERSPACE

Intelligence On, About and From the Net

"It's a ham-fisted attempt to keep a competitor out of the market. The guys down here on the plantation in Mississippi haven't figured out that indentured servitude has gone out of vogue."

— Royce Holland, former WorldCom Inc. vice chairman, on the Jackson, Miss.-based phone giant's legal attempts to prevent other former employees from working for him at Allegiance Telecom Inc., his new Dallas-based competitive local exchange carrier

THOSE RUMORS ABOUT MICROSOFT CORP. SELLING OFF THE MICROSOFT NETWORK just won't go away. Noise hears Microsoft (www.microsoft.com) may be contemplating a



face-saving solution, selling off a chunk of MSN to an outside partner while retaining a minority stake. The most legitimate partner? The usual parcel of telcos, dealers say. Problem is, most of them are too busy watching the next deal in the never-ending MCI Communications Corp. poker game to ante up. Anybody want in on a different game?

THERE ARE THREE WAVES OF PROFIT to be made on the Net, says Brewster Kahle, known for developing the Wide Area Information Service — known around cyberspace water-gate holes as WAIS. First, he says, you can write a book about the Net (been there, he says); second, sell a company — say WAIS Inc. — (done that, he adds). Finally, actually make

Noise

money on the Net. Kahle, a swindler that's coming real soon now from his new start-up, Alexa Internet (www.alexa.com), is the most legitimate partner? The usual parcel of telcos, dealers say. Problem is, most of them are too busy watching the next deal in the never-ending MCI Communications Corp. poker game to ante up. Anybody want in on a different game?







Join now **FREE!**

MENU

[Categories](#)

[Features](#)

[Weekly Picks](#)

[Browsers](#)

[Power Search](#)

RESOURCES

[FREE Newsletter](#)

[Help](#)

CNET SERVICES

[CNET.COM](#)

[NEWS.COM](#)

[BUILDER.COM](#)

[BROWSERS.COM](#)

[GAMECENTER.COM](#)

[SEARCH.COM](#)

[SHAREWARE.COM](#)

[ACTIVEVX.COM](#)

[MEDIADOME](#)

[SNAP! ONLINE](#)

MARKETPLACE

[BUYDIRECT.COM](#)

[CNET Store](#)

[How to advertise](#)

Featured Product

Advertisement

Alexa

(Alexa Internet)

Free Web Navigation Software

Why wander the Web aimlessly when you can use a guide? Get Alexa and always know where to go on the Web.

Alexa is a toolbar application that runs with your browser (including Netscape browsers and Microsoft Internet Explorer) to provide information on sites that you visit as well as suggested links on where to go next. These recommendations are based on the surfing paths of other Alexans, so you have tens of thousands of people helping you find the best sites on the Web. Alexa is the first anonymous, multisurfing, Web site suggestion tool and conferencing application.

Alexa also lets you send messages to other Alexans via its Instant Messaging service. It also has a desktop reference feature that always allows you to look at something else without leaving the page you're on. You can even review old Web sites (especially helpful when you get a "404 not found" error message) through the Alexa Archive.

Why bump around without a guide? Get Alexa and know where to go.

Please visit our website: www.alexas.com

Download

☒ [Windows 95/NT](#)

4

search
engines
?

John Markoff (left): deer caught in headlights.



Alexa Party

Text and photos by Tish Williams

We thought we'd lost Brewster Kahle under a mountain of 1s and 0s in his crusade to archive the entire Web. Miraculously, Kahle broke free, surfacing unsquashed on July 17 to launch a brand-new company, Alexa Internet.

Alexa offers a new type of browser that provides a synopsis of each site you happen upon, lets you know who owns the site, tells you how popular that site is based on other visitors' time there and recommends alternative sites of the same genre that are the most trafficked. Keeping Kahle's pet project on top of the Web, Alexa will share the fruits of its labor with the Internet Archive.

The martini-bolstered crowd responded enthusiastically when Kahle showed how Alexa draws from the archive during searches: If users try to click through to an item during their search and find that the page no longer exists on its server—resulting in the dreaded 404-Not Found error—Alexa displays a clickable icon that accesses an archived version of the lost page. "These are just some of the fun things you can do if you have five T3s in your backyard," Kahle said.

Imagine a world without 404-Not Found. Now there's paradise. Or an office with five T3s ... now *that's* paradise.



Stewart Brand (far right) and friends lap up free drinks.



Alexa CEO Brewster Kahle (center) gives questionnaires the lowdown.



Your Complete Guide to Searching the Net

- ◀ Introduction
- ◀ Editors' Choice
- ◀ Suitability to Task
- ◀ Summary of Features
- ◀ Labs Tests
- ◀ How We Tested
- ◀ Readers' Picks

Reviews

- ◀ Search Indexes and Directories
- ◀ People Finders
- ◀ Business Finders
- ◀ Usenet Search
- ◀ Metasearch



Your Complete Guide to Searching the Net

The Future of Search

These pages offer a good guide to searching, but it's apparent that such tools, no matter how advanced, still can't keep up with the Internet. The amount of data, the rapidity of change, and the hectic web of links pulling it together make truly coherent, comprehensive organization nearly impossible. People differ on where the problem lies--user interface, data collection and analysis, speed limitations of hardware, and so forth--but clearly, we need a new solution. Here is a sampling of new approaches from the information industry.

Mapping Companies such as Perspecta (www.perspecta.com) and Semio Corp. (www.semio.com) offer Java-based products that analyze and organize documents by concept and attribute. The products then respond to mouse- or text-based input with dynamically generated visual, navigable maps of relationships and hierarchies.

Collaborative Filtering Alexa Internet's (www.alexa.com) free downloadable toolbar stays on your desktop while you surf and provides statistics and owner information for the sites you visit. Using data on where other users who have visited a site have gone as well as link and text analyses of the site, Alexa dynamically suggests other links.

Specialization NewsBot (www.newsbot.com), from HotBot, is a standalone ActiveX control that you access from your desktop. You can search the NewsBot database of top news sites using an interface similar to a Web search screen.

Client-Side Metasearching Prompt Software's WebSleuth (www.promptsoftware.com) is another in the vast array of client-side search tools. WebSleuth lets you query an unlimited number of search sites simultaneously; the program analyzes returns and drops unrelated and broken links before giving you your results.

Personal Agents Inquisit (www.inquisit.com) targets business professionals with a subscription-based "personal intelligence service." Users set up agents with ongoing queries; the agents monitor Inquisit's database of news and information services and send e-mail updates at times specified by the user.

Human Contribution Sites like LookSmart and Yahoo! use people to analyze and categorize the sites in their databases. The Mining Company (www.miningco.com) takes this a step further with its cooperative of sorts. Users apply to be Guides for subsites on specific topics. Guides are responsible for the focus of their sites and for updating links and adding new ones.

◀ PREVIOUS

NEXT ▶

From the December 2, 1997 issue of *PC Magazine*

RELATED STORIES

- [ZDNet Products Channel: Web Search User](#)
- [ZDNet News: Internet search indexes](#)
- [NetBuyer: Internet/Online Software](#)

TOP

— A ZDNet Site —					
ZDNET HOME	SITE GUIDE	FIND IT	COMMUNITY	AD INFO	CONTACT US
Copyright (c) 1997 Ziff-Davis Inc.					



Cyber Historian

*As the Internet
grows in importance
and sheer size,
Brewster Kahle
proposes a
strategy to
archive the Net*

As the inventor of the WAIS (Wide Area Information Servers) system and founder of WAIS Inc., Brewster Kahle pioneered publishing on the Net. Kahle's latest project, Alexa, is an Internet archive – a time capsule of sorts – in an effort to preserve everything on the Net as history in the making.

CMQ: Why do we need to preserve a snapshot of the World Wide Web?

BK: The real opportunity in archiving and datamining the WWW is to produce a wonderful navigation service that gets better as it grows in content and users.

Archiving the Net is a necessary first step, but then making it useful to everyone is what the Alexa project is all about. Named after the Library of Alexandria, Alexa is looking to build a worldwide resource for those that use the WWW.

Alexa is a free service and is available from www.alexa.com. It will improve as more people use it because it learns from people. In many ways, we are building a huge digital library that becomes smarter as the Net grows and people use the service.

What progress have you made to date?

We have collected a full snapshot of the public Web, including pictures and video clips. Furthermore, we have kept collecting the changes. We have seven terabytes of data now, and it is growing at one terabyte a month.

To put this in perspective, all the text in the US Library of Congress is 20 terabytes. (A terabyte is a million megabytes, and a book is about one megabyte of text.)

The data has many uses and we are using it in many interesting ways including:

- Finding patterns in the Web to build a navigation service that learns from people. This is the core of Alexa's service. A lot can be learnt from "datamining" the link structure and aggregate usage patterns to guide future users of the Web. Search engines and directories don't scale up well as the Web grows.
- Tracking changes in the Web and its use to find areas of growing and shrinking interest: What's hot? What's up-and-coming?
- Offering historians and scholars an unprecedented collection of human voices.
- Keeping a record of the birth of a new medium, and the dreams people have for it.

What guidelines would differentiate a digital library from current directories?

Alexa is different from current directories and subject listings in the following ways:

- Alexa is a navigation service that uses the link structure of the Net, content of the Web pages, and the aggregated usage paths to suggest Web pages and sites that might be interesting to a user.
- Alexa also keeps copies of what was on the Web in the past to serve out-of-print Web pages.
- Alexa is not a keyword search system; it uses the paths people have made through the Web to suggest similar Web sites.
- Alexa is not a directory (like Yahoo) where a few editors are creating subject listings. It uses what many people have done to reinforce clusters that make sense.

Who would administer a global digital library?

Everyone. This is our Web. Everyone has a role to play in reinforcing the paths to the quality materials. Alexa serves as a conduit and a repository for usage paths and content to be used by everyone. We expect every company, government, and university to build systems that are similar in many ways to what Alexa is building.

What criteria would determine the kind of content admitted?

We gather all publicly accessible Web pages and revisit them every six weeks. We do not gather materials that people have indicated they don't want gathered by using the "robot exclusion standard" of the Net.

How would changing file formats ensure the usability of digital data over a period of decades?

Many obscure formats will not be readable, but the more popular ones will be because people want to see them again. We are trying to help by keeping a large collection of materials accessible so that it is worth writing converters.

Can digital archives spawn other technologies and services?

Imagine being able to see a video of your great grandfather. Imagine not panicking when your computer disk crashes. Imagine being able to do research on mountains of newspapers in a flash.

We are only starting to understand what can happen when computers and global knowledge get intermingled.

How will digital libraries seek to answer important questions about our culture?

As our cultural artifacts go digital (music, writings, scholarship, letters, ...), then the digital libraries will be a place where these are organised and served from.

We are looking for ideas and interested people to use the system and to suggest where it should go. <<





General Editorial Liaison:

Over the November-December period, the following journalists were also contacted regarding general Alexa inquiries, new angle pitches, feature article opportunities and follow up. These journalists have responded by either writing about Alexa or expressing interest in covering Alexa in the near future:

Sari Kalin CIO Magazine	Howard Millman InfoWorld	David Lidsky PC Magazine
Simson Garfinkel Freelance	Tim Jackson London Financial Times	
Gene Kaprowski Forbes ASAP	Elizabeth Wasserman San Jose Mercury News	

Compilation of Alexa Press Coverage

Headline:	Publication:	Date:	Author:
AIM's Who's News	AIM Newsletter (Association for Interactive Media)	12/27	Staff
The 7-Terabyte Man Takes on the Storage Gods	InternetWeek	12/8	Rivka Tadjer
If Links Could Talk	Release 1.0	12/97	Jerry Michalski
Missing in Action	California Real Estate Magazine	12/97	Bits & Bytes section
Head to Head: Internet Search Utilities	Windows Magazine	12/97	Amy Helen Johnson
Navigazione guidata	Plug (Italy)	11/21	no byline
From Philanthropy to Entrepreneurship	London Financial Times	11/25	Tim Jackson
Video cards put new life in your PC	Anderson Independent-Mail (Anderson, SC)	11/16	James Derk
Cool Tool of the Day	Cool dot Com	11/12	Dr. Vinton Cool
MyPoints	Advertising Age	11/10	no byline
You can nuke more bad guys with 3-D video card	The Sun Herald (Gulfport, MS)	11/9	James Derk
Video cards are better and	The Decatur Daily	11/6	James Derk

Headline:	Publication:	Date:	Author:
faster	(Decatur, AL)		
Bigger, better, faster with new video cards	Press-Republican (Plattsburgh, NY)	11/5	James Derk
Looks good: See what a new video card can do for computer	Leader-Telegram (Eau Claire, WI)	11/5	James Derk
The Godfather	Wired Magazine	11/97	G. Pascal Zachary
Alexa Internet	Signal Magazine (Fairfax, VA)	11/97	no byline
fyi	Windows Magazine	11/97	Staff
Site Reviews: Go There Find That	Yahoo! Internet Life	11/97	Staff
Alexa Party	Upside	11/97	Tish Williams

INTERNETWEEK

www.internetwk.com

The 7-Terabyte Man Takes On The Storage Gods

BREWSTER KAHLE HAS CHALLENGED THE STORAGE gods. Kahle's company, start-up Alexa Internet, has downloaded and stored no less than the entire contents of the public Internet, a staggering 7 terabytes and counting, as the basis for its new information business.

Kahle has developed a new Internet search engine/gopher technology that promises to deliver more reliable Web searches. His goal: Be the site with the most complete database so Web surfers always get the information they're looking for. No more "404 Error: URL Not Found" hits that pop up all too often. Kahle hopes revenue will be generated by advertising, and the search engine will be free to surfers, much like Yahoo! and Lycos, or any other search engine.

Of course, the 37-year-old president of Alexa Internet admits that few companies have such massive storage needs or goals. Nevertheless, he has seen the future and warns that the Internet will quickly have an impact on their requirements. "Fortune 500 companies now have the number of nodes on their intranets that the Internet had five years ago, and it's growing all the time," says Kahle.

High Performance

Kahle's storage M.O. is to buy as many high-speed disks as he can afford, which he claims access data at 9 megabytes per second, and then use DLT tape for the rest.

DLT costs about \$10 to \$12 per gigabyte instead of the \$100 per gigabyte price for hard disk. Of course, tape performance varies. If you're accessing data that's on a tape where the robotic arm is already residing, Kahle says you'll get a retrieve time of about 3 megs per second. If you have to change tapes, though, it can take as long as four minutes for the same amount of data.

"I use the tape to store data on Internet sites that people don't visit very often—like DOS community discussion sites," says Kahle. "For the really popular site data, such as search engines like Yahoo!, the information has to be on our hard disks. People can surf to our Web site and test our new gopher. If it takes people up to four minutes to find Yahoo! on our system, they'll go away and never come back."

Right now 1 terabyte of Kahle's data is stored on high-speed disks from Quantum Corp. and Seagate. He chose Quantum first because

it used to have the biggest hard disks Kahle could find on the market—9 gigs. "But now we have these brand-new 23-gig Seagate drives, and they're working well," he says. "And I'm a good test if something works well because we are so hard on the technology—constantly accessing. I think both of these disk drives are bulletproof."

Kahle says the drives are built well because they're intended for the PC and video industries. "Even if you don't go with Quantum or Seagate, go with a brand catering to these same industries and you'll have a fairly safe bet," he says.

He also buys Quantum's DLT tape technology, and then uses hierarchical storage management software from Large Scale Configuration Inc. (www.lsci.com). The LSC software determines what should be stored on tape and what should be on hard disk for fast and constant access. Kahle says LSC comes from the mainframe world and understands the issues surrounding converting a centralized storage system into a cost-effective multitiered one.

Full-Time Service

"What the Internet brings to the table for us—and many other companies with both Internet and intranet Web storefronts—is that we're open to provide customer service and information 24 hours a day, seven days a week," Kahle says. "That means access to stored data has to be absolutely optimized for what people want, balanced with what you can

afford."

Smart storage management is critical to Kahle, who says if the data now becoming available on corporate intranets is to be of real use, it must be stored and accessed in a meaningful way.

For Kahle, real use means people in organizations can learn from one another to share knowledge across departments. And Kahle insists that type of storage management won't break your budget.

"Try what we're doing on the Internet on your intranet—I promise it won't cost a million dollars," says Kahle. "Go for the \$100-per-gig hard disks where you can, and in many cases you may be able to store the whole thing on hard disk, at least for now. The average Web site on an intranet or the Internet is measured in megabytes, so even 100 gigs worth of storage is \$10,000." •



Alexa's Brewster Kahle to Fortune 1000s: Plan to spend on storage.

Tim Jackson • On the Web

Archive holds wealth of data



America is full of entrepreneurs who spend their careers making money and their retirements disbursing it to good causes. Cases where the business comes after the philanthropy are rarer, but that is the story of Alexa Internet (www.alexa.com), a company that has developed a useful tool for navigating the web.

Alexa is the creation of Brewster Kahle, who worked on fast computers at a company called Thinking Machines. He founded a company called WAIS, which he sold to AOL in 1995 for \$15m (£8.8m).

Too young to retire, Mr Kahle started musing on the fact that the web is transient, because no trace remains of a web page once its publisher has removed it. With his own money, he built the Internet Archive (www.archive.org), which now contains eight terabytes of data. Since a terabyte is 1,024 megabytes, the archive's holdings are equivalent to about 16m full-length books.

The data are stored on

tapes in a juke-box system, which allows any one of the world's 135m web pages to be retrieved in about 15 minutes. The archive is updated by software spiders, which crawl around the web looking for new pages and adding them to the archive.

While spending \$100,000 of his own money building this system, Mr Kahle realised the benefits of developing a navigation tool to help people find what was most useful on the web. He chose the name Alexa in homage to the great library at Alexandria, one of the wonders of the ancient world.

Alexa is a software "agent" in the form of a small toolbar beneath the browser window. The program sends information on your behaviour to Alexa which can be used to benefit others.

The toolbar keeps track of which page is being displayed in the web browser window and offers information related to that page. By clicking one button on the Alexa toolbar, you can answer the question "Where am I?". A window appears showing the popularity of the site, details of the web site's owner and statistics on the site's updates.

The toolbar also suggests where you should go next, using a number of methods. One is to track usage paths. The program keeps track of the sites you visit in each session and how long you spend there, and relays this information to Alexa. The software gives greater weight to pages that are viewed for longer, on the plausible grounds that they are likely to be more useful.

Another approach is link analysis. The Alexa spiders find sites that recommend the page you are looking at, and considers other sites to which they link as possibly related. A third approach is content analysis.

Using these methods, the Alexa toolbar creates four links to sites that may be related to the current page.

Even though the Alexa client activates six processes on the company's server every time the browser shows a new page, Mr Kahle says the computing power needed for the continuous updates is modest.

Mr Kahle hopes to turn this into a lucrative advertising vehicle. From the company's point of view, its key advantage is that the toolbar stays open for hours as users surf the web, allowing Alexa

to sell far more ad "impressions" than even the most popular web site.

Alexa can "slice and dice" its customer base because it asks users for demographic data such as residence, age, job and sex. Advertisers can buy toolbar ads that are shown only when the user is looking at a particular web site. This lets Alexa sell site-specific ads without the site owner's involvement.

So far, Mr Kahle and his colleagues have not sold a single ad. They want to avoid putting users off by making the ads too obtrusive at the beginning.

But there is another reason for their hesitation. Two months since the product's launch, the software has been downloaded by 75,000 people. Microsoft and Netscape often achieve similar download numbers in a matter of days.

Mr Kahle says his download target is 1m copies by the end of 1998.

If Alexa cannot achieve enough users, others will imitate or incorporate its best ideas. The company needs some marketing tools to encourage other web sites to promote its product.

tim.jackson@pobox.com

Review of Alexa Internet Software

These are my impressions of Alexa version 1.2. If you have an earlier version, you should upgrade to version 1.2. If you have version 1.2, future updates should be automatic, according to Alexa. I am assuming that you have already read the overview, quick tour and FAQ on the Alexa web site. They are well-written and will answer many questions about Alexa's features. If you are the sort of person who usually ignores documentation, you should make an exception in Alexa's case. A [collection of articles about Alexa](#) is available.

I downloaded and installed Alexa with no problems.

Alexa worked with Netscape 3 and 4 and Microsoft Internet Explorer 3 and 4.

Alexa places an icon (a lower-case "a" in a blue circle) in your taskbar's tray--the tray is on the far right side of the taskbar, next to the display of the current time. If you have closed Alexa and want to re-open it, click on the icon.

The Alexa toolbar can be dragged around the screen--I prefer putting it between the bottom of the browser window and the Windows 95 task bar. There is an Autohide option that will keep the toolbar hidden until you want it, thereby giving your browser window more room on the screen. If you use large buttons in your browser toolbar (I use small text-only buttons in Netscape) and you use 640x480 screen resolution, you may find that you must use Autohide--otherwise, if you have the Alexa toolbar displayed along with all the other bars on the page and you are viewing a web page with frames, you may have little room left in your browser window!

Help is launched by clicking on the extremely tiny question mark at the lower right side of the Alexa toolbar. Help can also be launched by clicking on the lower case "a" on the left side of the Alexa toolbar, which reveals a menu that includes help, technical support Q&A and other options.

The "Where You Are" menu gave me a lot of information that I usually didn't need to know. Maybe others will find it to be more valuable. The "View Site Map" option on the "Where You Are" menu is useful if you want a "road map" of what a site looks like; however, it doesn't give you every page on a server--it only gives you some of the pages that branch off from a domain's main page. Alexa uses the [SiteMap](#) site to provide site maps; SiteMap has some [interesting features](#).

I wonder how many Alexa users will bother to use the "Vote" button. Unless people use it, the like/dislike ratings won't be meaningful. A drawback of the "Vote" feature is that the votes will be irrelevant in some cases. If someone goes to Geocities and doesn't like a member's page, it will count as a "dislike" vote for Geocities as a whole, but that vote has no relevance to any of the other Geocities member pages.

Alexa has small ads on the toolbar.

The "Where To Go Next" list has up to four choices on the toolbar--clicking the right-pointing circled arrow just to the right of the "vote" button gives more choices. "Where To Go Next" usually gave useful suggestions, but occasionally the suggestions seemed worthless or inappropriate. For example, when I went to a page on the www.sierra.net server, Alexa

suggested a bunch of links to Sierra Leone. Alexa's suggestions will get better as more people use Alexa and Alexa gathers more data.

The archive icon (it looks like a pillared building) turns into a little piece of paper if you go to a web page for which Alexa has an archived copy. If you go to a page that gives you a "Not Found" message, such as "File Not Found", "404 Not Found" or "HTTP/1.0 404 Object Not Found", wait a couple of seconds to see whether the little piece of paper appears. If it does, you can click on the archive icon and Alexa will fetch its archived version of the page. A page will appear that says:

Your page is being
fetched from our Archive.

You will be notified by Alexa
when it is ready to be viewed.

Because the data must be fetched from our tape
robot, this process takes time. The normal wait for a
page and all of its images is about 15 minutes, but
complicated pages may take considerably longer.

You may continue with your normal use of your
computer while the page is being retrieved, but please
leave Alexa running so that it can notify you when your
page is ready.

When Alexa has retrieved an archived page, a window appears on your screen asking you whether you want to look at the archived page.

To test the Alexa archive, I looked through my bookmarks for "Not Found" pages and I also used AltaVista's advanced search function to find other peoples' bookmark pages that hadn't been updated since early 1996. Most of the "Not Found" pages I encountered were in the archive. This may be partly due to the fact that some of those "Not Found" pages were personal pages that were probably not on the web for very long or were seldom visited while they were on the web. If you are looking for a "Not Found" page that was once popular, it is probably more likely that it will be in Alexa's archive.

An archived page retrieved by Alexa doesn't have the same URL as the "Not Found" page. For example, when I asked Alexa for the following "Not Found" page:

<http://www.hyperion.com/~koreth/particles/>

the URL of the archived page was

http://widener.alexia.com/http/www.hyperion.com/80/~koreth/particles/_ja_index.html

A day later, I was able to retrieve the above page immediately by using the longer URL. This was not because my browser had cached the page. It was undoubtedly available because Alexa had cached it. An Alexa technical support person told me that the Alexa cache holds many gigabytes of frequently requested data.

Alexa gives you the most recently archived version of an archived page. This is not always useful, because conscientious webmasters often put up a placeholder before they remove a page completely. For example, Alexa retrieved one archived page that had only the following text:

This link is gone!
Please remove bookmarks to it.

If Alexa could give me an earlier archived version of that page, it would have been more useful to me. (On the other hand, if the webmaster put up a link to the new location before the page was taken down, the archived page would link to its new location.)

As another test, I tried to find an old Simpsons page. The people behind the TV show "The Simpsons" have threatened legal action against some people who created Simpsons pages. I found a "Not Found" Simpsons page that appeared to be archived, but I got the following message from Alexa:

The object you requested is not in the Archive.
We will not be able to fulfill your
request.

The probable cause for this error is that the
authors of the page requested its removal
from the Archive.

Such requests are honored immediately, but
there is a delay before all databases are
updated to reflect the deletion. So Alexa will
continue to indicate the availability of the
page, for some time after its removal.

In a framed site, Alexa won't tell you whether a particular frame is archived unless you open the frame as a separate page.

It is not possible to retrieve an archived page that is from a domain that no longer exists, such as the once-popular www.gnn.com. You can't get a "Not Found" message because the server itself isn't connected to the Internet any more. An Alexa technical support person said they are working on a method to permit users to access a page from a domain that no longer exists, but it is not currently possible.

The EB icon launches a window that allows you to search a dictionary, a thesaurus, and a concise version of the Encyclopedia Britannica. The full Encyclopedia Britannica is also an option in that window, but you may use the full Encyclopedia Britannica only if you have paid the Encyclopedia Britannica subscription fee.

Alexa has the most astonishingly responsive technical support I have ever seen. They have replied to my questions the same day and the replies have not been form letters. I wish other software and hardware companies had tech support like that!

A friend of mine told me he installed Alexa but he doesn't use it. He said he didn't like the clutter on his screen and he thought it was like Pointcast in the sense that it was software that initially seemed interesting but later became an annoyance. He admitted that he hadn't read the documentation and that he had been unaware of some of Alexa's features.

Send questions and comments to scruffy@prairienet.org. If I made any mistakes in this page, please help me correct them.

[scruffy home](#)



Web applications.



FILEMINE

Search for files:

Win (all)

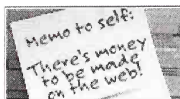
Search

[Advanced Search](#)[Your Download List](#)[Site FAQ](#)[Jewels](#)[Packs](#)[Comparative Digs](#)[E-mail Newsletter](#)[Home](#)[Categories](#)

Sponsored by:

Take an e-business
test drive...

CMPnet RESOURCES

[Home](#)[Site Map](#)[Search](#)[Ad Info](#)

Jewels

Great Finds

PC Mac

Alexa 1.2

[Download this file now](#)

With IE 4.0 in wide release, many users have gotten a glimpse at Microsoft's chosen way to bring the Internet to your desktop. What the company forgot, in my opinion, is some of the substance serious surfers need to complete their browsing experience. Alexa Internet is a free service that fills the gap, helping you surf smarter, not harder.

Overall Rating:

★★★★★

Licensing: Freeware

Requirements: Win 95 or NT; 32-bit Winsock connection

Publisher: [Alexa](#)

Internet

Download size: 1,658 KB

Add file to
download list

Alexa displays a toolbar that works with your browser, providing a continuous source of relevant recommendations on where your next stop on the Web should be. This toolbar is displayed over your browser's status bar. Information in the Alexa bar is site- and context-sensitive, changing as you surf from site to site. Available information for a given site includes who owns the site, how many pages it has, how frequently the site is updated, and the site's popularity among other Alexa users. The latter information is provided by giving users the ability to vote on whether they like or dislike a particular site. Alexa also provides relative, targeted recommendations of where to go next.


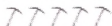
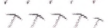
Using Alexa's network, you can see which of your friends are online and send them messages in real time, perform research on the Web through your complete desktop reference tool, and access the Encyclopaedia Britannica Online and Merriam-Webster's Dictionary and Thesaurus.

Alexa offers some very valuable information, although you may be overwhelmed at first with its abundant set of features. Considering the improvements that have been made to this app since we first featured a review of the beta version, you can look for more good things from Alexa in the future. --
[Christopher Spera](#) and [Kerry R. Krueger](#)

Ratings Criteria

Ease of Use:

★★★★★

Interface: 
Documentation: 
Performance: 

CMPnet SPECIAL REPORTS



NetBusiness offers
Windows 95
commerce software.
Download it [free](#).



Listen Up: Audio on
[building small-biz
sites](#) with Windows
95.

Today's ► [Business Guides](#): For the low-down on high-tech companies.
Features ► Stop wasting time with failed downloads. Use [crash-proof FTP](#).



© 1997 CMP Media Inc.



business video

[click here for the latest in business video](#)

THE WALL STREET JOURNAL INTERACTIVE EDITION.

SMALL BUSINESS SUITE

December 4, 1997



FRONT SECTION

MARKETPLACE

MONEY & INVESTING

TECH CENTER

SPORTS

PERSONAL JOURNAL

- NEWS
- FAVORITES
- PORTFOLIO

EVERY PAGE
EVER PRINTED

NATIONAL
GEOGRAPHIC
INTERACTIVE

ALL ON CD-ROM
Check it out!

In this Section:

[Marketing/Media](#)

[Health & Science](#)

[Business Focus](#)

[Law](#)

[Who's News](#)

[Small Business](#)

[Columnists](#)

You Can See
EVERY PAGE!
Ever Printed!

NATIONAL
GEOGRAPHIC
INTERACTIVE

ALL ON
CD-ROM

SHOW
ME!

[Table of Contents](#)

[Related Sites:](#)

Small Firms Turn to Web To Boost Market Presence

By RIVKA TADJER

Special to THE WALL STREET JOURNAL INTERACTIVE EDITION

What's in a brand name? For companies seeking to turn a profit, everything.

But it's also expensive to transform a company name into a household word. So small businesses with invariably tight advertising and marketing budgets are turning to the Web for a unique opportunity to leverage bigger companies' market presence by buying ad banner space on their sites. Call brand-building on the Web the business version of social climbing.

Links

▶ Join an on-line discussion about using the Web for brand-building.

Wharton's Pennsylvania Small Business Development Centers
www.libertynet.org/pasbdc/

Alexa Internet
www.alexa.com

Cirque Corp.
www.cirque.com

Kraft Foods
www.kraftfoods.com

Tripod
www.tripod.com

iVillage
www.ivillage.com

"Striking strategic alliances on the Web and trying to leverage the brands of bigger entities makes a lot of sense; that's why everyone is trying to get in with America Online or put a banner ad on Yahoo! or Barnes & Noble," says Paul Morin, director of the Wharton Small Business Development Center. "Remember that the only way to take a small business out of obscurity and into profitability is by building a brand name. Today that means advertising in the big four -- TV, radio, print and on Web communities."

While TV, radio and print communicate with the consumer in only one direction, the Web allows for an interactive exchange. The approach companies take to building a brand on the Web should take advantage of that difference, Mr. Morin says. "This means exploiting what the Web medium can accommodate -- use audio, video, interactive e-mail, whatever it takes for people to buy into whatever you're selling and come back to your

[Barron's Online](#)
[SmartMoney](#)
[Interactive](#)
[Careers.wsj.com](#)
[Business Directory](#)
[Publications Library](#)

Search/Archives:

[Search](#)
[Briefing Books](#)
[Quotes](#)
[Past Editions](#)
[Journal Links](#)
[Special Reports](#)

Resources:

[Help](#)
[New Features](#)
[Your Account](#)
[Contact Us](#)
[Glossary](#)

Advertising:

[Advertisers](#)
[E-Mart](#)

company."

What the Web does best, Mr. Morin says, is reach a target demographic of males from the ages of 15 to 30, and, to a lesser degree, 30- to 40-year-olds. "These are the people using the Net all the time," he says. "And while there are an increasing number of women, target your interactive campaigns to get as many of this demographic profile involved with your company as possible."

Some brand-building veterans say the brand name is even more important than the product. Small businesses cooking up Web strategies may do well to take a page from this Madison Avenue school of thought.

"Your brand name defines the relationship consumers have to your company, not just your product. If that brand name is good and the relationship a loyal one, you can build your business with a variety of successful products and services," says Thomas Quarton, president of Cirque Corp., the Salt Lake City-based start-up company that invented the technology for [Apple Computer's](#) Powerbook touch-pad pointing device (called the GlidePoint), which replaced the trackball. In the five years since Cirque was founded, it has taken command of almost 50% of the touch-pad marketplace.

In 17 years of consumer marketing, Mr. Quarton has built brand names for Clearasil, M&M Mars, Marriott Corp., and Vidal Sassoon. But getting the word out on Cirque is the first time Mr. Quarton has felt how much of a difference the Internet makes. "People buy based on the relevance your company has to their lives overall, not just because of one product or another. That's what brand name really is," he says. "The Internet lets us directly interact with consumers -- showing demos of our product, amusing people with games on our Web site, having conversations over e-mail, being there live for them in as many ways as we can invent," he says.

"You want to know my favorite Web site? Kraft Foods," says Mr. Quarton. "I think it does the best job of inviting consumers to have a relationship with the company, not just one product Kraft sells. My wife and I constantly go into their recipe section and find great stuff. Kraft has pulled off the tough task of creating its own interactive community."

Since most smaller businesses don't have the funding to create their own communities, they are advertising on the sites of existing communities to build their own brand names. Banner advertisements on computer-community Web sites such as ZDNet are central to Cirque's Web campaign because they target Cirque's audience in an efficient way, Mr. Quarton says. "We advertise in a bunch of the computer print magazines that Ziff publishes, such as PC Magazine and Family PC -- a lion's share of our \$1 million print ad budget is for Ziff -- but our \$15,000 banner ad on ZDNet reaches readers of all those magazines at once." ZDNet's fee buys Cirque one fiscal quarter of advertising. Mr. Quarton also places a banner ad -- at the same price -- on the Yahoo! search engine site to reach less tech-focused Netizens.

If a business can only afford to do one brand-building program, Mr.

Quarton says, demonstrating prowess of the product at the point of sale is the way to go. "And if your product or service can be demonstrated and sold on the Internet, it may be the least expensive, most efficient way you'll find."

Netizens, Mr. Quarton says, are a curious and willing universe of prospects, and businesses can find out more on-line about what those prospects think than is possible from TV or an ad in a magazine. A registration form Netizens fill out for a contest can be a well of information. And the contest is good public relations for the whole company -- not just a single product.

Bruce Gilliat, general manager of a San Francisco start-up called Alexa Internet, says brand-building on the Web offers another advantage to folks in his boat -- the "25 employees but no revenues yet" situation: "We started our company and Web site in July, and even though we didn't have money for banner ads, we could swap ads with other companies on the barter system to create a virtual community," he says.

Companies such as Williamston, Mass.-based Web developer Tripod and Ivillage: The Women's Network swapped ads with Alexa.

Alexa's flagship product is an Internet search service. Alexa staffers have downloaded the contents of the public Web into a database and analyzed the sites. They keep updating their enormous database all the time. Owner Brewster Kahle claims that the free Alexa service can find any site on the Web, even if other search engines fail at the task and simply pop up with one of those pesky "Error 404: URL Not Found" messages. Alexa's slogan is "Where am I and where do I want to go next?"

Mr. Gilliat says all of Alexa's brand-building focuses on boosting the company's reputation, not just the Web-search product slogan. With \$1.3 million in seed capital from investors, the few paid marketing moves Alexa has made so far include spending \$100,000 on a launch party in the Bay Area, holding a press conference, and participating in an Internet trade show in Chicago.

Mr. Gilliat says the company also spent close to \$20,000 for a designer to make sure that everything -- the Web site, the banner ads, the brochures, everything with the Alexa name on it -- has precisely the same look. "That unified look is key to brand-name recognition," says Mr. Gilliat.

Next on Alexa's agenda -- when the company starts generating revenue from advertisers -- is probably TV advertising on cable stations such as CNNfn and CNBC. "And all the ads will have our URL and slogan at the end."

Businesses trying to tie TV, print and radio commercials back to the Internet should use the URL as part of the brand character itself, Wharton's Mr. Morin says. "Generally, when you go for branding, go for the community with the largest target audience based on what you can afford," he says.

The trick with Web-community deals, though, is avoiding being

taken advantage of by the big brand. "They hold the cards, because there are only so many big companies on the Web, and they have a huge universe of small companies to choose from," Mr. Morin says.

But it is possible to succeed in deals with large companies. Long Island, N.Y.-based 1-800 Flowers cut a deal with AOL earlier this year that was expensive and negotiated without any standard measure. It took a reported \$25 million for 1-800-Flowers to talk AOL into making the florist the only florist in AOL's popular Shopping Channel for the next four years. There are no standardized prices for these sorts of deals, AOL officials say.

Donna Lucolona, director of interactive services division at 1-800-Flowers, says there's a better way to gauge value than standard pricing: "If you can make back the investment in the first year, the price doesn't matter," she says. If the only florist the 10 million AOL account holders see every time they enter the Shopping Channel on AOL is 1-800-Flowers, Ms. Lucolona says her brand name goals have been accomplished.

[Return to top of page](#)

Copyright © 1997 Dow Jones & Company, Inc. All Rights Reserved.



UPSIDE.COM

[HOME](#) | [ABOUT UPSIDE](#) | [NEWS](#) | [LINKS](#) | [SEARCH](#) | [STOCKS](#)

Interview: Yahoo's Tim Koogle

December 01, 1997
By Richard L. Brandt

Yahoo just may be the most visited site on the Internet. Every day, perhaps a million people view some 50 million pages on Yahoo, making it one of the few profitable Web sites on the Net--from ad sales, no less! CEO Tim Koogle tells how it's done.

SIDEBARS

Interview:
[CNET's Halsey Minor](#)

[CNET](#)
[Company](#)
[Details](#)

[Yahoo](#)
[Company](#)
[Details](#)

UPSIDE: What brings so many people to your site?

Koogle: A number of things. The original premise was that of bringing functionality to the Web, of providing a navigational guide to a broad array of users. We wanted to be the only place you had to remember to come back to, to find access to all kinds of content, and now all kinds of goods and services as well. We aggregate content from different sites. Doing that created a platform. That platform ends up looking like a gateway, or a hub, so now tens of millions of people come through that hub monthly.

How do you know what people want? We use the consumption patterns of our visitors to augment and extend the platform globally.

But you use human brains, rather than technology, to choose new links? Yeah. It's important to separate the different means of content delivery from the structure of our sites.

Full-text search engines are an automated way of guiding people toward content. People enter a word or a phrase and hit "search," and a search engine goes and searches across a compressed index of text documents, which have been gathered by a crawler [an automatic software engine that gathers information about sites across the Web]. That's one way of connecting people with content.

Another way, which was our starting point, was to build a directory, a hierarchy where we aggregate sites around subject themes. We build a table of contents, if you will, and allow users to visually browse through that and be

guided by it.

We've done both. The way we figure out what people want in deeper aggregation and the way we structure our properties is part automated and part editorial. We have a lot of tools that we've built in-house to take all the massive usage files that we generate daily and reduce that and extract the fundamental key metrics. We also have internal tools that allow us to make use of that and deliver ads on a targeted basis.

But when it comes right down to it, people who have the sensibilities of what is the most logical, easy, fun, quick way of guiding people toward content make those decisions.

We seem to be so fascinated with technology, smart agents, automatic search engines, natural language programs and other artificial intelligence programs, but we forget that stuff doesn't work well yet. That's right.

What do you think of Brewster Kahle's approach [at Alexa Internet] of trying to lead people to new sites that may be of interest to them by showing them where other users go after visiting the same site? That's one of a number of different efforts that various folks on the Web have mounted, trying to detect user patterns dynamically and use that to either automatically or non-automatically suggest to them places they should go. This whole business of collaborative filtering and the rest has been tried, and none of it has worked super-well. As you know, we've had a relationship with [intelligent-agent company] Firefly, and we've experimented with collaborative filtering.

The question is whether users would take the time to enter in some of their preferences, to kick-start these things. Then whether, after using them for a while, the users perceive enough value in them. Thirdly, whether users have some sense that they may be missing something because something else is making the choice for them. We've gotten input from almost all the users in our focus group studies that there's a lingering concern that it excludes things they may want.

There's still a technology problem? I think so. It's too early.

Network Boss?

When the Web first arrived and first became a publishing medium, we all knew it would not be like television, newspapers or magazines—it would be its own form. Is Yahoo becoming that new medium? Yahoo is not becoming the medium. The Internet is the medium. I have always thought that the Internet was a

fundamental shift in the way people are able to access content and communicate with one another. You can boil that all down into a communication platform, but it's one in which people communicate and share content with one another. That content could be as simple as an e-mail or more detailed author content.

What Yahoo has been from day one is a media company, in the context of the Internet as a medium.

If the Internet is the medium like television, you're like one of the broadcast stations. Yeah, those analogies have been drawn before--that what we are building is a successful broadcast-media company.

Is your goal to be the primary starting place for people going onto the Internet, or one of a few big network companies from which people would choose? There's more than enough room for several entry points onto the Web. In fact, the Web is less controlled than any broadcast medium of the past, where people could buy a spectrum in a given geographical region and basically control the onramps into their broadcast medium. There are hundreds of thousands, if not millions, of entry points onto the Web. We have always intended to remain a hub or a gateway.

The online world started with services such as CompuServe and America Online, which offer content on their proprietary networks. But once the Web took off, these companies started being displaced by the Web. Are you, rather than MSN [Microsoft Network], the Web replacement for AOL? Some people are saying that. We live our days here at Yahoo trying to figure out how to satisfy our users' needs, to deliver as comprehensive a set of content as we can possibly build for as wide a customer base as possible globally, and add communications and community services. The fact that those same things apply to a well-done proprietary online service leads people to try to draw direct analogies between us.

America Online has been creating more of its own in-house content. Is that also an appropriate direction to take for a company such as Yahoo? We don't intend to do that. There's plenty of good-quality content on the Web.

Can proprietary online systems survive? Yes. There is a segment of the world's population whose needs are well-served by a proprietary service. There is a level of safety and ease and stability. Whether it's large enough to build a big business around, with the right economics, remains to be seen. It requires getting to a higher level of scale because you carry a higher cost base.

No Spam!

What's your position on exclusive partnerships with content creators? In general, we don't believe in exclusive relationships. People have asked us why we don't require exclusives from content vendors. In general, single-customer businesses don't work. For us to require exclusivity by any of our vendors means we would heavily limit their ability to establish and maintain a viable business. I won't do that.

On the inward-bound side of things, sometimes content suppliers or commerce partners request exclusives with us. It's tricky because we have a clear position that we don't limit the availability of content.

We do limited-term things, things we call "runway projects," where we carve off an initial period during which we don't work with many others [to launch the project]. But then that expires.

GTE recently sued Yahoo and five Baby Bells, claiming you offer exclusive access to their Yellow Page services, which cuts out GTE's own competing Yellow Pages. Isn't that an exclusive relationship? What's going on there? I'll say two things about that. First, GTE was offered the ability to be there on the Net and turned it down. So their position puzzles us. The second is that users are not at all excluded from getting to them. You can get to GTE's Yellow Pages from a number of points in Yahoo and on Netscape [NetSearch]. Go figure.

You also have relationships with Compaq and Gateway, so that a Yahoo icon appears on the desktop within Microsoft's Internet Explorer. Do you have any kind of exclusivity in that? Can they put a search engine or another navigation site on their desktops as well? There are levels of exclusivity. There are different parts to our relationships with them. We are the exclusive featured directory provider, but they can have other navigation companies cofeatured.

So while you don't go after exclusivity, you don't mind trying to negotiate yourself a premier status. You bet.

It seems as though there is also some value to users to have you make an editorial judgment about which service or content is better. To some extent you have to do that, don't you? Yeah. We're careful about those sorts of things, though, because our users vote and show us what they find most compelling.

How do you strike a balance between users' choices and your choices? When you go to www.yahoo.com and go through the subject branches of the tree, you get to one with a listing of sites. You'll always find that those are alphabetically arranged. We want to stay unabridged and unbiased. You also see the "cool" sunglasses by some sites, indicating the picks of the day. These are editorial

judgments.

You have done a great job creating the Yahoo brand. Some people say your brand mainly appeals to people first venturing onto the Internet. Do you see yourself as a starting place for beginners? We're more broad-based, and the data shows it. About 60 percent, and sometimes more, of our traffic comes from people accessing Yahoo from the workplace. We also have survey data done by third parties that tells us that well over a third, probably about a half, of our users are experienced users who have been on the Web a long time.

You've recently added things such as free e-mail. Given that that's not likely to be a great revenue prospect, what do you gain from offering that to users? It's a key component to this thing called communication and community. We're not only a place where people come to get information, we're a place where people can find other people. It gives our users yet another level of quality experience.

Another thing it does for us is create "stickiness" to Yahoo's properties. It gives us yet another form of direct one-to-one communication with our user base because they register for the service. People also register for [My Yahoo](#) and [Yahoo Chat](#). It allows them to tell us what they want. I'm also fundamentally convinced that it turns into revenue.

Because it keeps people on your site, viewing your ads more frequently? Definitely.

What's your philosophy on creating stickiness? There's still a common mistake among Web sites, trying to make it so you never leave that site, which is the antithesis of the whole spirit of the Internet. There's a balance in there. What we don't do here is try to figure out ways of tricking people to stay on Yahoo. We have, probably at the top of our values, a commitment to make it easy for users to find and get connected to content, no matter where it is on the Web.

However, we want to continue to extend our value to users. As we do that, we find that users come back and stay longer. It's a subtle thing, but it's profound.

Are people reluctant to register? What do you do with the information? We never give it to anyone else. We've been asked to sell it many times, but it stays inside Yahoo, and we use it only to the extent that it helps us to guide the delivery of content to them. Ultimately, we may begin to target ads to them as they consume inventory. With the recent addition of free e-mail on Yahoo, we've made a pact with our users to not spam them.

Future Technologies

You have a relationship with Visa. Do you see a future as an electronic-commerce site? Yeah, with one subtle distinction. Just as we don't intend to get into the original-content business but partner with original-content creators, we also don't intend to get into the business of literally selling things on Yahoo. We do intend to be a good aggregation source for various vendors of goods and services. We've got a short list of a dozen to 15, going on 20, commerce partnerships: E-Trade, [Ameritrade], DLJ Direct. Some of them we haven't announced.

So you'll point people to these sites and work out some revenue-sharing deals, take transaction fees? Both. We end up being a piece of virtual shelf space, if you will. We'll get a fixed fee for the distribution space, then a cut of the transactions.

What about MasterCard and American Express? We have agreed with Visa to have them as the premier card, but the others are not excluded.

You are also doing audio broadcasts. May I presume at some point video broadcasts as well? If streamed video emerges, then we will take the same position of aggregating directly to the source. We'll offer people a convenient place to come and find [video] content and then deliver it to them.

What do you think about push technology? As you know, there's a range of technologies and forms for push.

Some people will have some of their content delivered some of the time via push. I've always believed that you'll see a spectrum. There are people on one end of the spectrum whose content needs are narrow and don't change often. On the other end are people whose content needs are broad and change all the time.

The people on the narrow end of the spectrum can more easily designate stuff to be delivered to them regularly via push. The folks on the other end will have a hard time designating their needs.

We intend to embrace as broad a user base as possible, so push is a viable component. But it is not the be-all and end-all.

You've also done some deals with WebTV and Microsoft. What do you think about the merger of the television and the Internet? There's something to the broad thesis. There will be a convergence between what we call a separate computing device and a television. Televisions will get smarter and bandwidth will increase, but it will take time. Several generations of consumers now have been conditioned to view and use the TV as a one-way entertainment device.

So it requires not only a huge change in technology, but a behavioral change for people. Yes.

Do you think that will happen? It'll happen over time. I can easily see five, 10 years from now that we'll have in our homes and our workplaces devices that have multiple uses. Cost will continue to plummet until it enables the convergence of multiple functionalities for similar or lower price points than we see now in PCs. That makes sense. Less clutter in your life.

Why not just two separate devices, one for push and one interactive? Possibly. The way we view it, it's all good. The more content that's available on more devices, the more consumers there are accessing it, and the more there is for the role of Yahoo in guiding them to it.

Everybody's Favorite Nemesis

What do you think about sites such as MSN, which act as aggregator and original-content provider? Everyone who hangs out with the folks in Redmond knows that they're evolving their strategy and their vision. They started off with a strict approach of being a proprietary online service and charging a subscription fee and offering access. But they're searching for ways of evolving out of that.

Will Microsoft ever be a significant competitor to Yahoo? Microsoft remains a company that we admire a lot and fear like everybody else! We have a relationship with them, and we continue to look for great ways of augmenting that, and they are in the process of continuing to evolve their vision. We're mindful that they could come into our space.

What worries you most? One thing that is a challenge for anybody in our space is managing our growth. It turns out to be something we've done fine on to date, but it remains a concern, and it probably always should be until the day I set foot in my grave!

Richard L. Brandt is editor of UPSIDE.

[Home](#) | [About UPSIDE](#) | [News](#) | [Books](#) | [Events](#) | [Links](#) | [Stocks](#) | [Search](#)

Feel free to send e-mail to: [Ad Sales](#), [Editorial](#), [Subscriptions/Back Issues](#), or [Feedback](#).
[Subscribe to UPSIDE today!](#) Copyright ©1997 Upside Media Inc. All rights reserved.
Powered by [Thunderstone](#).

internet.com

internet
news.com

Business News

E-Commerce News

Finance News

International News

Intranet News

ISP News

Product News

WebDeveloper
NewsInternet Stock
Report▶ Internet Advertising
Report

Mecklermedia

▶ Internet.com

Magazines

Internet World
Trade Shows

Advertise

Corporate
Information

Search

Subscribe

This is your world.

internet advertising report

Alexa to Offer "Ads on the Bar"

[December 11] Free Web navigation service Alexa said that its advertisers now have more opportunities to reach target markets by displaying ads directly on the Alexa toolbar.

The service, which gives suggested links about where to go next on the Web and detailed information about every Web site, calls the program its "Ad on the Bar" campaign.

The free Alexa service is displayed as a toolbar on the user's desktop and is available for users of Windows 95 and NT. Alexa can be downloaded from the company's Web site.

Designed to make Web advertising easier and more focused because Alexa accompanies users while they surf, Alexa's toolbar displays advertising images on a continual basis as they surf the Web. Advertisers can target users by interest area, industry, demographics, and by advertising on their competitors' home pages.

The company said that "only Alexa enables advertisers to advertise directly on competitor sites. For instance, Barnes & Noble can display its banner ad every time the user visits the Amazon.com site, as well as other competitor or high- traffic sites."

Initial advertisers include CMP, WhoWhere, CNET and Web Turbo. Rate card data was not disclosed.

IBM®

Help for the Holidays

Bookserve

GO

**StorageTek**

Click for FREE Stress Wolf.

Stressed over information storage?**TECHWEB**

Search CMPnet:

[Search](#)
[Advanced Search](#)**Sections**[News](#)[Stocks/Finance](#)[Web Development](#)[Product Reviews](#)[Company Profiles](#)[Trade Shows](#)[ActiveX/Java](#)[Careers](#)[Opinion](#)[Humor](#)**Support**[Encyclopedia](#)[Feedback](#)[E-mail Newsletters](#)**Featured Sites**[CMPtv](#)[Net Insider](#)[NetGuide](#)[Shareware](#)[Games](#)[Horoscopes](#)**CMPnet RESOURCES**[Home](#)[Site Map](#)[Search](#)[Buy Software](#)[Buy Books](#)[Ad Info](#)[Sponsor Index](#)[Affiliates](#)**Tech Search**[Search Home](#)[Advanced Search](#)[Search Help](#)[Search the Web](#)[Direct Article Link](#) | [New Search](#) | [Search the Web](#)

November 01, 1997, Issue: 811

Section: News/New Products

**fyi**

Where am I? How'd I get here? Where should I go next? If you've ever asked yourself these questions while surfing the Web, now you can find your path with **Alexa** Internet. The free Web-navigation service works with Web browsers to provide a toolbar at the bottom of your screen for a continuous source of relevant recommendations of where to go on the Web. The recommendations are derived by analyzing patterns and anonymous usage paths. You can find the service at <http://www.alexa.com>.

Copyright (c) 1997 CMP Media Inc.

[New Search](#) | [Search the Web](#)

You can reach this article directly here:

<http://www.techweb.com/se/directlink.cgi?WIN19971101S0069>

integration of Encyclopedia Britannica's reference materials coincided with the launch of **Alexa** 1.0, the company's first release of the service since the beta program was launched in July 1997.

Owing to the growing popularity of the **Internet**, this type of service will undoubtedly provide valuable information to both users and advertisers on the World Wide Web. It will provide helpful hints to users as to the most effective means for navigating the Web and inform Web site owners of the popularity of their sites.

Request further information on the following related reports: Frost & Sullivan's Report # 2880-63, U.S. Internet Service Markets, and Report # 2613-63, Consumer Internet Fee-Based Services.

IRG # 63; A/T Code 0811

© 1997 Frost & Sullivan. All Rights Reserved.

[Telecom Headlines](#) | [Telecom Archives](#) | [Telecom Home Page](#)
[Newsletter Search](#) | [Contact Us](#) | [HOME](#)

Search text: **Alexa Internet**
1/100

Previous Document	Search Results	Next Document
-------------------	-----------------------	---------------



The
GIF ANIMATION
WebSite



HOME PAGE
WEB DESIGN
AUTHORING
CLIP ART
ANIMATION
WEBTOOLS
BROWSERS
WINDOWS
BUSINESS
FINANCE
EDUCATION
HEALTH
NEWS
GAMES
MUSIC
TRAVEL
RECIPES
AWARDS
ADVERTISING
FEEDBACK



Web Tools

CLIENT SIDE

[SERVER SIDE](#) [CLIENT SIDE](#)

Alexa

Alexa has come of age, and we rate it as one of the most valuable free resources for WWW surfers. Alexa uses an humungeous archive of WWW sites. It provides powerful search capabilities, technical and ownership information on sites, and as a bonus allows immediate free access to the entire Encyclopaedia Britannica and Merriam-Webster's Dictionary & Thesaurus



Anawave Websnake

Anawave WebSnake 1.23 will grab, search, and retrieve files from the Internet. WebSnake allows users to filter unwanted data and retrieve the information they want. WebSnake supports off-line browsing, advanced searching and retrieval of files by keywords or file type, website mirroring, creating graphical site maps, and retrieval of e-mail addresses. WebSnake also supports the ftp. This is powerful software, almost essential for anyone who who does lots of surf research. Windows95 & NT.



AutoWinNet95

Automated Internet for Windows95/NT. Allows scheduling common tasks for unattended operation. Upgrades automatically, and supports 35 steps, including FTP: retries busy sites, wildcard download and upload, make/remove remote files and directories, plus customized logins. Sends email with an advanced, feature packed editor, unlimited multiple mailboxes, checks mail, or cleans up your mailbox. WWW: Grab

Alexa

Op zoek naar een slimme zoekmachine? Probeer Alexa. Alexa zoekt gericht, presenteert concurrerende websites van jouw favorieten, vertelt je of een website snel en populair is en weet ook nog heel oude webpagina's boven water te halen. En dat is nog lang niet alles.

Het bestbewaarde geheim van het Internet?

Vlak bij de beroemde Golden Gate Bridge in San Francisco ligt het Presidio, een door hoge sparren omgeven legerkamp dat ongeveer honderd jaar geleden werd opgericht. Militairen wonen er niet meer, het grootste deel van het kamp is museum geworden. Hoewel we er voor het eerst zijn, komt het kamp ons bekend voor. Naar onze stellige overtuiging moet hier een van Columbus' beste afleveringen zijn gefilmd. In een van de witte barakken van het kamp is Alexa Internet gevestigd, het best bewaarde geheim van het Presidio en van het World Wide Web.

Alexa werd in 1996 opgericht door Brewster Kahle en Bruce Gilliat. Bekende namen voor Internet die-hards, want beiden waren in het verleden betrokken bij WAIS, een zoekstelsel voor gespecialiseerde tekstbestanden op het Internet. WAIS, voortgekomen uit het bedrijf Thinking Machines, was er al voor (inmiddels archaïsche) navigatiehulpmiddelen als Gopher, en dus ook ver voordat het World Wide Web zijn intrede deed. In de meeste Internet handboeken wordt nauwelijks nog aan WAIS gerefereerd, want de meeste WAIS-servers zijn inmiddels 'normale' zoeksystemen geworden. Het gelijknamige bedrijf werd in 1995 voor vijftien miljoen dollar aan online gigant America Online verkocht.

Alleen nuttige info a.u.b.

Met Alexa borduren Brewster Kahle en Bruce Gilliat voort op de WAIS-gedachte. Namelijk hoe vind je in het doolhof van het Internet relevante informatie over onderwerpen die je interesseren? Reguliere zoeksystemen overladden je met duizenden suggesties, en wegwijzers als Yahoo doen dat ook. "Theoretisch kun je natuurlijk het complete web indexeren", zegt Brewster Kahle in zijn kantoor.

"Maar wat moet je met al die informatie? Het web verdubbelt zich iedere zes maanden, voor mensen is het vrijwel onbegonnen werk om orde in de chaos te brengen." Een wegwijzer als Yahoo of Excite verwijst misschien naar slechts een procent van het web, aldus Kahle. "Zelfs als je met veel freelancers werkt, zoals The Mining Company (zie kader), krijg je hooguit honderdduizend pagina's boven tafel."

Alexa gebruikt drie geautomatiseerde technieken om interessante webpagina's op te sporen. Met behulp van een gratis programma kunnen webgebruikers aangeven of zij bepaalde sites nuttig vinden of niet. Ook volgt Alexa de bewegingen van gebruikers op het World Wide Web. Daarnaast wordt rekening gehouden met zogenaamde verwijspagina's van experts. Kahle: "Op het web vind je altijd wel een paar goeroes die weten waar Abraham de mosterd haalt."

JAN LIEBENGA

In feite maakt Alexa gebruik van wat Kahle noemt 'het collectieve geheugen van de webgemeenschap'. "Ik huur graag speelfilms waarin bepaalde acteurs spelen. Die namen kan ik natuurlijk in een filmencyclopedie opzoeken, maar ik heb veel meer aan de aanbevelingen van mijn filmvrienden. Alexa werkt net zo."

Informatie over informatie

Tot zover de theorie, nu de praktijk. Om te beginnen moet je bij Alexa (www.alexa.com) een gratis programma ophalen, circa 1,7 megabyte groot, dat onder Windows 95 en in de toekomst ook onder MacOS en Windows 3.x als een taakbalk op je scherm verschijnt. Met dit programma kun je via zogenaamde popup-menu's extra informatie over webpagina's opvragen. Alexa noemt dat meta-informatie, ofwel 'informatie over informatie'. Zo kun je bijvoorbeeld achterhalen wie de eigenaar of exploitant is van een site. Gaat het om een beursgenoteerd bedrijf, dan wordt niet alleen het adres, maar ook de omzet vermeld. Een deel van die informatie komt van The Securities and Exchange Commission (SEC), de waakhond van de Amerikaanse beurs, en InterNIC, dat domeinnamen registreert en opslaat in een database. Maar er is meer: Alexa vertelt je ook of de webserver door-

gaans snel of juist erg traag is, uit hoeveel pagina's de complete site bestaat en hoe vaak die pagina's worden geactualiseerd. Verder staat er een waarderingscijfer bij. "Stel dat je via het Internet bij een reisbureau een vlucht naar Bali wilt boeken", zegt Kahle. "Dan wil je wel wat meer informatie over dat bedrijf hebben. Met Alexa weet je in elk geval of de site wordt gewaardeerd."

Alexa wil de dienstverlening overigens uitbreiden. Zo is er een samenwerking aangekondigd met TRUSTe, een organisatie die er op toeziet dat bij elektronische transacties de privacy wordt gewaarborgd. Ook werkt Alexa samen met The Recreational Software Advisory Council, die websites voorziet van 'ratings' voor taalgebruik, sex en geweld. Daarnaast is er een koppeling mogelijk met de Encyclopaedia Britannica, een thesaurus en een woordenboek (Merriam-Webster Online's WWWebster Dictionary). Als je een woord tegenkomt dat je niet kent, kun je snel het synoniem opzoeken. Via pop-up-menu's wordt ook naar andere relevante sites verwezen. Als je bij een pagina van een reisbureau bent beland, zal Alexa zijn concurrenten noemen.

404 Not Found bestaat niet meer

Uniek is de mogelijkheid om van iedere webpagina de vorige versie van die pagina op te vragen, bijvoorbeeld als je een '404 Not Found'-mededeling (pagina niet gevonden) te zien krijgt. Alexa archiveert sinds 1996 het complete World Wide Web. Het beheer daarvan is uitbesteed aan de non profit organisatie Internet Archive (www.archive.org). Op dit moment bestaat het archief uit vijfhonderd tapedrives, elk met een capaciteit van ongeveer twee terabyte. Om het complete web te archiveren is vermoedelijk ongeveer tien terabyte nodig. Ter vergelijking: 's werelds grootste bibliotheek, het Library of Congress, heeft een omvang van twintig terabyte.

Het Internet Archief is volgens Kahle bittere noodzaak. Kranten en tijdschriften worden bewaard, maar webpagina's niet. Onderzoek van de Universiteit van Colorado heeft uitgegeven dat de levensduur van een webpagina hooguit 44 dagen bedraagt. Wie wil weten wat er op 5 november 1995 zoal op het Internet gebeurde, kan nu in elk geval bij het Internet Archief terecht. Samen met het Smithsonian Institute werkt het Internet Archive aan een tentoonstelling over de invloed die het web had op de presidentiële verkiezingen van 1996. Op dit moment kan

Op zoek naar kwaliteit

Hoe vind je kwaliteit op het web? Hoe scheid je het kaf van het koren? Menige webgebruiker zal het onmiddellijk beamen: gemakkelijk is het niet. De meeste webgebruikers beginnen bij elektronische wegwijsers als Yahoo, LookSmart en Excite. Deze zoeksystemen worden samengesteld door redacteurs die iedere dag het web afstruinen op zoek naar relevante informatie. Maar deze redacteurs zijn omnivoren, en geen specialisten.

Argus Clearinghouse (www.clearinghouse.net), een initiatief van de Universiteit van Michigan, verwijst wel naar allerlei specialisten op het web.

Het verst daarin gaat **The Mining Company** (www.miningco.com). Dit Amerikaanse bedrijf huurt freelancers in die lijsten met relevante websites op hun vakgebied samenstellen. Zo wordt de pagina over kindergeneeskunde onderhouden door een echte kinderarts. The Mining Company kent als hoofdcategoryën onder meer kunst en amusement, zaken, carrière en onderwijs, gezondheid, sport en reizen. Iedere gids moet een korte opleiding volgen waarin hij of zij vertrouwd wordt gemaakt met webtechnieken als HTML en Java. De eerste maanden wordt hun werk doorlopend geëvalueerd. Als iemand toch niet de gewenste kwaliteit levert, wordt hij uit zijn functie gezet. De freelancers krijgen veertig procent van de bruto advertentie-inkomsten als honorarium uitbetaald. Inmiddels telt The Mining Company ongeveer duizend gidsen. Binnen vijf jaar moet bij The Mining Company over ongeveer vierduizend onderwerpen informatie kunnen worden opgevraagd.

Edison (www.edison.com) probeert ook orde in de chaos te scheppen via een overzichtelijke webpagina met honderden verwijzingen en een index van zoekmotoren. Als je wilt kun je de complete site downloaden en offline bekijken.

CompuServe (www.compuserve.nl) is eind vorig jaar Communities gestart. Via zeven hoofdbrubrieken belandt de gebruiker in submenu's met een verdere onderverdeling. Ook toegankelijk voor niet-leden. Zie ook het artikel op pagina 60.

het archief overigens alleen met behulp van het archief-programma worden ontsloten, in de toekomst wellicht ook via zoeksystemen. Commerciële exploitatie is vanwege het auteursrecht vrijwel uitgesloten. En: "Als mensen echt niet willen dat we hun pagina's bewaren, halen we ze onmiddellijk uit het archief", zegt Kahle. "De meeste mensen komen echter al gauw op hun besluit terug." Zelf verdient Alexa hoofdzakelijk aan advertenties die in de taalkalk worden weergegeven. Kahle: "We kijken ook naar andere distributiemogelijkheden van onze technologie. De bedoeling is dat iedereen straks weet wat Alexa Internet is."

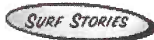
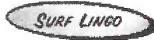
UR's

www.alexas.com
www.archive.org
www.clearinghouse.net
www.miningco.com
www.edison.com
www.compuserve.nl



You published all **50** of marketing's color data sheets onto the company website.

YOUR GUIDE TO THE BEST OF THE WEB



Q: Guru,

I've been looking for a browser that can analyze Web sites when I surf to them -- you know, like who owns them and how much traffic they get.

A: While the new browsers have more bells than the cathedral at Notre Dame and more whistles than NFL referees, they still don't have what you're looking for just yet. But a piece of Web navigation software called Alexa does.



Once you download and install the free (but advertising-supported) software, Alexa shows up as a small toolbar that hovers over your browser. The "Where You Are" section of the toolbar provides a mother lode of info on any site you connect to, including who owns the site, even providing a phone number and street address for many (this feature is limited to sites with .com, .org, .edu, and .net suffixes in their names, however). Alexa will reveal a site's loading speed, total number of pages, and when it's been updated. It also gives you an idea of how much traffic it gets -- if you're planning on forging a mutual link with someone else's site, this is the perfect tool to help you determine if it's worth your while.

The "Where To Go Next" section of the toolbar suggests links to other sites that it thinks are related to the page you are visiting. These links haven't been suggested willy-nilly -- they come from Alexa watching how people move between sites, then analyzing these movements to determine which sites are the most helpful and logical places to surf

next.

But for me, the best part of Alexa is its ability to raise the dead. Dead sites, that is.

Did you know that the average life span of a Web page is, according to Alexa, 44 days? That's shorter than an NBC midseason replacement sitcom. Try accessing a site that's gone belly-up, and you get that annoying "404 Not Found" message. End of story, right? Not anymore. Alexa archives a vast amount of the Web on its central servers -- when you run across a dead site, you can connect with its service through the toolbar and ask it to grab its copy of that site. Currently, Alexa stores about 8 to 10 terabytes of such information. For perspective, bear in mind that the data in the Library of Congress would equal about 20 terabytes in ASCII form.

Alexa works with Netscape Navigator 2.0 or higher, Netscape Communicator 4.0., and Microsoft Internet Explorer 3.0 or higher. Currently it only runs on Win95 or Windows NT 4.0, but a Mac version is coming soon.

ARCHIVE CATEGORY: [Communicating](#)

[Today's Question](#)

[Surf Guru Archive](#)

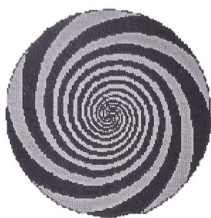
[Yahoo! Internet Life Daily](#)

ASK THE SURF GURU: FILL IN YOUR QUESTION BELOW:

Your e-mail address (so we can contact you, we won't post it):

Your name (so we can say who the question's from):

Your question to the Surf Guru:



hypnosis.com

Our site went down in December (understatement, the server is still being repaired). To make matters worse our back up failed to recover anything. Up until Friday Jan 18th, we were running on a two year old back up. Then a customer e-mailed us and said we could find our site at www.alexacom.com well, needless to say we are back up because of them. If you have a web site make sure you have alexa back you up.



A multi-award winner, the Alexa software is both free and extremely useful. It works more or less as a browser's torch-light.

This is a little program that will add a horizontal bar to the bottom part of your browser. And as you browse, it points the way, and will inform you of the popularity of the site you currently visiting without any bias, as well as other useful information. All on real time and hassle free.

Eurosat.com is one of the first satellite TV website that has been rated **Top 10000**

The raw data is collected from all the major backbone around the world, hence far more reliable than other dodgy survey services, Such as Top XXXX etc. Results are generally very accurate. Exception is when certain website carries far too many tiny pictures on each page, throughout the entire site. As each little GIF or JPG translate into one hit, it may ultimately effects the total outcome. But even this event rarely occur.

If in doubt, compares the rating to **Alexa visits** for consistency. Generally speaking,
Top 10 is reserved for the Goddess in the world such as AOL and Microsoft.
Top 100 is for the Internet Royal Families (CNN, abcnews),
Top 1000 is for The Rich and Influential (Demon, Euronet, Techweb).
Top 10000 for the bests in its field,
Top 25000 for excellent well design websites,
Top 100000 for websites that begins to show its potential.

So what are you waiting for? [Download it now.](#) Their website is waiting for you.

If you are getting a Low Traffic (meaning less than Top 100000), there is no reason to be panicked. You do only if you have already invested millions and years.

This web site does not offer pirate Sky card nor Sky-related software
This site is in no way related to **Eurosat Distribution(UK)** or **Eurosat magazine(Italy)**
All enquiries regarding the web site should be directed at The Webmaster
© defiant@eurosat.com 1997. All Rights Reserved. DISCLAIMER

The Eleventh Annual MicroTimes 100

As we went from conference to conference in 1997, we encountered the same phenomenon over and over. ✧ Internet startups founded in 1995 described themselves as "old." And most of them that were still alive in 1997 had changed their business model—several times. Despite these unpromising conditions, some of them were actually starting to make money. ✧ At the same time, erstwhile titans stumbled—aside from Apple's well-publicized woes, CompuServe was absorbed into AOL and WorldCom, Informix and Sybase hit walls in the enterprise market, Silicon Graphics saw its traditional markets eroded and its brains drained by startups, and so on. ✧ The titans that didn't stumble, led by Microsoft and Intel, exhibited a phenomenal lack of complacency. ✧ All this being so, it will probably come as little surprise that there are a lot of people on this year's list you've probably never heard of—founders of startups, one-person businesses, intrepid tech-support guys, and relative unknowns who somehow make the industry a little easier to deal with. Some of them will no doubt vanish into peaceful obscurity by next year; some will probably live to be Bill Gates' biggest headache (or next acquisition). ✧ Also prominent on the list are some folks we haven't seen for a while—in many cases because they've been building their next startup after the last one gave up the ghost, and the new one is doing something pretty interesting. ✧ We'd like to thank our readers once again for bringing many of these deserving souls to our attention. We also encourage you to send us email at editors@microtimes.com to reproach us about the equally deserving souls we may have missed!

The Eleventh Annual MicroTimes 100

continued from page 126

for major development strides by emerging Russian software development companies. His own company produces award-winning software. But, his major accomplishments are centered on promoting other small Russian developers.

"With letters of recommendation from Microsoft and Intel, with whom he works closely, he brings companies into cooperative agreements to help them polish their business practices and products for global distribution. His efforts help qualified individuals produce an honorable living in Russia's moribund economic atmosphere. He is building an entire business culture, promoting production, innovation and growth. His company provides an avenue to the Western world.

"Anet Software is one of the most established development companies in Russia. Mr. Gorjushin built this company with virtually no outside investment and no personal experience during historic turmoil and transition while Russia changed from Communist to a Democracy. He survived through the tumultuous fall of Gorbachev's Soviet Union and the failure of Yeltsin's Federation of former Communist Republics. Anet thrives now, in the constantly changing rules of the new Russia."

We're happy to join in saluting Mr. Gorjushin, and highly recommend visits to his Web site.

Andy Grove
Chairman and CEO
Intel

www.intel.com

Frankly, we're not sure about all those dancing engineers in neon space suits, but they're an unforgettable icon for Grove's latest market push into the world of video, from conferencing to gaming. That in turn is only the latest indication of his tireless quest to anticipate and capitalize on emerging markets. Meanwhile, back in the mainstream, he also keeps the Pentium family well ahead of its competition as the silicon platform of choice.

Jeff Hawkins
Chief Technical Officer
Palm Computing

<http://palm.pilot3com.com/>

After years of struggling with the assorted problems of pen-based computing, something that's been occupying him since the GridPAD,



Hawkins and his team finally hit the sweet spot with the Palm Pilot. The little PDA was a blockbuster in 1997, bringing new credibility to a market that's been longer on hype than functionality, and spawning an entire industry of applications and add-ons.

Frank Huang
Chairman and CEO
UMAX Data Systems
www.supermac.com

While Huang has long kept Taiwan-based UMAX a key player in the arena of PCs, motherboards, scanners, and other devices, his most stunning accomplishment during 1997 was retaining UMAX/SuperMac's status as the only worldwide licensee of the MacOS at a time when Mac clones were experiencing a Jobs-induced bloodbath. Not only is UMAX offering its customers OS 8-based models, Huang extracted a pledge from Apple to certify future UMAX products and include MacOS 8 in all new UMAX system designs (excluding those based on CHRP) through July, 1998.

Echoing the sentiments of a grateful populace, Huang commented: "Mac clones today remain an important part of a healthy Macintosh market, and UMAX is committed to leading this vital market segment with world-class products and world-class partners such as Apple. UMAX has always had a strong relationship with Apple. We are confident that we will maintain this relationship and continue to contribute to the advancement of the Macintosh platform, and expand the Macintosh market as a whole."

Dave Hughes
Managing Partner
Old Colorado City Communications
www.oldcolo.com

Unlicensed-spectrum wireless's biggest evangelist, Hughes made excellent progress during 1997 with his National Science Foundation grant to deploy wireless Internet access in Colorado's San Luis Valley. The service is up and running, delighting rural schools and discomfiting the local telcos prepared to charge through the nose for similar access.

International Brotherhood of Teamsters
www.teamster.org

Normally the Teamsters are a bit far afield from the high-tech industry, but their successful strike against UPS helped focus attention on a similar inequity in the treatment of the computer industry's contract workers. We hope the cautionary episode resonates in the proper quarters.



Irwin Jacobs
Chairman and CEO
Harvey White
President
Qualcomm

www.qualcomm.com

Qualcomm's co-founders have built the San Diego company into a leader in wire-

less communications, from satellites to cell-phones—to say nothing of the various Eudora email packages. Departing a bit from the strictly technological, they also entered into a deal with the city of San Diego by which the facility housing Chargers and Padres games will henceforth be known as Qualcomm Stadium.

Steve Jobs
Acting CEO
Apple Computer
www.apple.com

"He has resurrected Apple from the brink of insignificance!"
enthuses one reader.

Well, maybe. It remains to be seen whether founder Jobs has succeeded in reversing Apple's death spiral, though the G3 machines represent a much-needed step in the right direction. We wish him well, particularly since a Mac-less world is too horrible to contemplate.



Brewster Kahle
President and CEO
Alexa Internet
www.alexa.com

Having sold WAIS to AOL, Kahle has found himself a new project—a Web navigation tool designed to learn from its users and produce affinity-based

recommendations of where to go next. The Alexa software is immeasurably aided in this endeavor by one of Kahle's ongoing projects—archiving the entire Internet.

Stephen Kahng
Founder and CEO
Power Computing
www.powercc.com

The unchallenged winner of the 1997 No Good Deed Goes Unpunished Award, Kahng effectively kept the Macintosh platform alive while

useit.com → [Alertbox](#) → Feb. 1998 Reputation management

Jakob Nielsen's Alertbox for February 8, 1998:

The Reputation Manager

There are two ways for users to know which of the two million websites to visit:

- **Brand:** the user knows that a site usually provides certain qualities; if the user likes these qualities, then it will probably be worth visiting more pages on that site
- **Reputation:** the user follows the advice of *other* users who know that a site has quality

Branding is the focus of most current Web projects: the theory is that building a powerful brand while the Web is still relatively small will allow a site to be profitable when the Web gets larger. A example of a good Web brand is **news.com**: when you want to know what's happening in the information industry, you will usually be able to find recent news at news.com.

Even though brands work well for a few, large sites, they are not a good mechanism to help users handle millions of sites. On the contrary, the nature of the Web encourages the formation of many smaller sites, and most of its value comes from such specialized sites. Thus, the Web needs a mechanism for making sense of **overwhelming diversity**.

Since there is no way for computers to automatically measure quality, we have to rely on human judgment for Web quality ratings. The reputation manager is a way to automate the processing of such human judgments; not a way to make the judgments themselves. In other words, **quality needs to become an explicit attribute** of Web objects.

My vision for a reputation manager involves the coordination of billions of individual quality judgments by hundreds of millions of users. Every time you encounter an information source on the Internet, your Web client software will present you with an opportunity to vote on its quality. Typically, this would be done by adding two buttons to the interface: a **thumbs-up** button and a **thumbs-down** button. A neutral rating would be given by *doing nothing* (since we want to minimize overhead in the user interface), but when a user encounters something particularly good, he or she would hit the "good" button. Similarly, disappointing services would be punished by a click on "bad."

The simplest reputation manager would compute the average rating for each information source, but more advanced services would use ideas from collaborative filtering and compute different ratings for different users. Basically, the reputation manager would find **other users whose tastes are very similar to your own** and give added weight to these users' ratings. Since the Web will have half a billion users in five years, it will always be possible to find other users who match your interests, no matter how obscure they are. Thus, the reputation manager can deal differently with people who love the Spice Girls and people who don't.

The reputation manager will collect ratings for entire websites, for individual pages, and for people who contribute comments to discussion groups or chat rooms. The resulting reputation can be used to direct users to sites that will be helpful or interesting, and it can be used to filter out the less valuable contributors in chat rooms and online discussions. In this way, the reputation manager becomes a more powerful version of the **bozo filter**.

Reputation management will be especially valuable when combined with micro-payments: once you have to pay for clicks, you will be motivated to find out *in advance* whether the destination website is any good.

Implementing Reputation on the Web

Initially, I expect reputation managers to become embedded in proxy servers for large corporations and offered as a value-added service by larger Internet Service Providers. In both cases, it is possible to collect data about the behavior of a large number of users at a single point. In about three years reputation management will become an Internet-wide service that users can subscribe to by paying a small micro-payment for each recommendation.

There is already an independent reputation manager available on the Internet: [Alexa](#) - which unfortunately is a browser add-on and thus not fully integrated with the user's Web client software (of course, right now, browsers serve as a weak type of Web access clients). The most relevant features of Alexa relative to this column are:

- **Reputation statistics** for most sites on the Web showing how frequently they are visited and how popular they are. For example, [Alexa's page about www.useit.com](#) shows how many of Alexa's users like my site and how many don't like it.
- **Recommendation links** to good other sites that are related to the current page.

Even without any fancy statistics, **most websites could benefit from explicit use of quality ratings** in their interface. A simple logfile analysis will show you what parts of your site are the most popular, and it would be reasonable to give these pages special prominence in search results and in listings. For example, my own [list of old Alertbox columns](#) highlights those columns that have attracted the largest readership in the past. Doing so prevents new readers from being overwhelmed by choices and allows them to focus on the links that are most likely to be of interest. A site with an even larger set of old material could provide a single [page listing nothing but its top hits](#).

February 22: Tracking the growth of a site

See Also: List of [other Alertbox columns](#)



Alexa

Brewster Kahle invented the WAIS (Wide Area Information Server) system and then went on to start WAIS Inc, an electronic publishing company. Bruce Gilliat was VP of sales and marketing at WAIS Inc.

So? Well, they're now the guys behind Alexa, a new Net service built on a question messrs Kahle and Gilliat asked themselves.

Q:"We all experience the gaps in navigating and finding information on the Internet as it is used today--frustrating keyword searches that turn up hundreds or thousands of web pages and sites, very few of which are of any interest. What if we, as a community of users, could effortlessly pool our collective experience and add human intelligence to navigation? What if we could fill in those gaps? It is a radical concept ..."

Their answer was, of course, to create the Alexa service, "the first Internet software product that learns from people".

And, "This intelligent navigation service provides a continuous source of relevant recommendations of where to go next on the Web," says the Kahle/Gilliat team. "Alexa works in parallel with, and independently of Web browsers, and is displayed as a toolbar on the user's desktop. The more people use Alexa, the more we all benefit."

Where's the benefit for K&G? In their bank balances, of course. But what else is new? Alexa provides context-sensitive advertising in its "Where to Go Next" selections and the blurb says, "While you will receive advertisements when using the service, the ads will be relevant to what you are viewing and *in some cases may act as additional 'Where to Go Next' selections.*"

In short, you could be forgiven for cynically thinking that Alexa *just* might be yet another push media app thinly disguised as a browser plugin. But for all of that, it's fairly cool, if only for its ability to give you an instant snap of sites you land on which subscribe to it. And even if they're not members, you still get a little info. That's us on the right ;) It also gives you a means of keeping in touch with other people who're hooked up to Alexa by allowing you to send and receiving messages.

Download Site: [Alexa v1.20](#)

PROGRAM BUILDERS:

If you'd like us to review your software, please contact [Jon](#) and we'll be glad to take a look.

Rating Cell:

Usefulness:




Performance:



User Friendly:



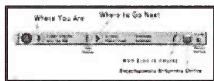
 *Fable Of Contents*

Manning the Wires



by Ric Manning

ALEXA BRINGS DEAD WEBS BACK TO LIFE



Brewster Kahle has found an answer to those pesky "404 File Not Found" error messages: He simply copies the entire World Wide Web to his hard drive. If someone decides to take a site or an individual page offline, he's still got a copy.

That might strike you as a fairly harebrained idea, but Kahle has been knocking around the Internet for quite a while and he swears he knows what he's doing.

In the early 1980s, Kahle developed WAIS (Wide Area Information Server), one of the first systems for publishing and indexing information on the Internet. WAIS eventually lost out to HTML, but America Online was impressed enough that it paid Kahle several million dollars for the WAIS technology. Kahle used the money to set up shop in San Francisco's Presidio and create a company and a product called *Alexa Internet* (www.alexa.com), a new web navigation system that Kahle named after the lost library of Alexandria.

Every few months, Alexa takes a snapshot of the public portion of the Web. By Kahle's count, that's about 640,000 web sites and more than 130 million pages, plus all the messages posted on Usenet newsgroups. The whole thing is stored on about 500 data tapes for a total of about 6 terabytes, about as much data as you'd find in a large city library.

With the whole Web on tape, Kahle said he is able to study it in ways that others can't. "We have a deep knowledge of what the Web is and we are trying to produce tools for people to make better use of the Net," he said.

The first of those tools is a browser add-on called Alexa that can be downloaded free from Alexa's web site. The program creates a toolbar that

helps web surfers understand where they are and where they might want to go. Bring up a web page with Alexa running and it will tell you who registered the domain name and the number of pages at the site. It will also rate a site's speed, quality and popularity.

"This is not a tool for someone who just wants to see pretty girls dancing across the screen," Kahle said. "We think of this as a really good tool for web masters and information professionals. It's for people who are going to a lot of different sites and trying to make sense of what they see. It will tell you who's behind this web site, how big is it and can I trust it?"

Kahle said the system uses InterNIC domain registration data to find a site's owners. Traffic estimates are based on counts from proxy caches at large network access points. Alexa's own crawlers measure server speed. The quality ratings are contributed by Alexa users.

The Alexa toolbar will also bring dead web page back to life. If you get a "404" error, you can hit Alexa's archive button to get the most recent version that Alexa captured. An example is the web site created by the Heaven's Gate cult. The pages are no longer online, but Alexa can serve them from its archive.

Alexa uses a robotic system to fetch tapes from its archive, so retrieval sometimes takes several minutes. But if Alexa has the page, it will display it with all accompanying graphics.

Kahle said Alexa copies only material that is offered to the public. The Alexa indexing agent bypasses pages that contain a robots.txt file, which tells web crawlers that they don't want to be indexed.

"We'll respect that," he said. "There's even an explanation of how to use the file on our site."

At the same time, web masters can invite Alexa to archive their sites by filling out a simple form on the Alexa web site.

Bringing web pages back from the dead is a pretty good trick, but it may not be Alexa's best. Using anonymous traffic data from some network access points, Alexa determines where surfers go when they leave a particular web site. Every time you visit a site, the Alexa toolbar tells you what paths were followed by the people who came before you.

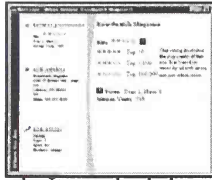
Check in at the White House site, for instance, and Alexa will might point you toward the sites for the House of Representatives, the U.S. Senate and THOMAS, the congressional information site.



suggest sites that don't have clear associations or sites that are related, but new to you.

"There are communities of people who circle around information," Kahle said. By tracking their surfing patterns, Alexa can identify groups with common interests.

Many of Alexa's links seem obvious. At *Boardwatch's* home page, Alexa points to three sites that deal with bulletin board systems. At www.msnbc.com, for instance, Alexa will point you to other TV network sites, to Comedy Central and MTV. At other times, Alexa will



And Kahle sees Alexa developing more services that could be useful to web masters. For instance, Alexa may be able to determine what kinds of people use a particular site.

"A lot of our users have given us demographic profiles," he said. That information can be used to match web sites with specific groups of people. It can also help people who use the Web have a better understanding of it.

"The search engines are going to the 10 channels," said Kahle. "We don't think that's what the Web is about. We think it's about diversity. People who use the Web go to a phenomenal number of servers. They're not just there to get the weather and the sports scores."


WEB LUBE

Could your site use a quick checkup? The web mechanics at the Web Site Garage (www.websitegarage.com) are only too happy to critique your work.

Type in your URL (or someone else's) and the garage will deliver a five-point report that gauges load time, looks for dead links and spelling errors, and evaluates your HTML design. It also calculates the popularity of your site by counting how many other web sites link to yours.

You can use the report to make improvements or you can hire the garage's team to give your site a makeover.

Editor: Jack Rickard - Volume XI: Issue 2 - ISSN:1054-2760 - February 1998
Copyright 1998 Jack Rickard - ALL RIGHTS RESERVED

 [Fable Of Contents](#)

WORLD

The Voice of Enterprise Computing

February 23, 1998

WINDOW MANAGER · BRIAN LIVINGSTON

Alexa provides additional ways to navigate the Web

Now a new, free service can give you vital information on how popular a Web site is and how many people liked and disliked their interactions with it. Even better, you can jump to related sites without having to resort to a keyword search, and you can retrieve old Web pages that have dis-



WINDOWS USERS ARE increasingly exploring the World Wide Web. Fewer than 20 percent of Windows PCs don't have Internet access, according to a recent Microsoft press briefing. But looking at a Web site and having confidence that the site is legitimate and trustworthy are two different things.

appeared from the Net.

Alexa is a relatively small free-ware download (918KB in size), that works with Netscape Navigator 3.0 or later and Internet Explorer 3.0 or later on Windows 95 and Windows NT. (If you are using Windows 95, you need a version of the file Kernel32.dll dated later than Feb. 2, 1996. You can download the updated version from Alexa's site.)

When you install Alexa, a small toolbar appears at the bottom of your screen when your browser is active. While you're viewing a Web site, clicking the Where Am I button on Alexa's toolbar reveals a wealth of information.

Leading the display is a ranking of the number of hits the site receives based on an analysis of traffic across the Internet backbone.

Even more useful is the name, address, and phone number of the site's owner.

This information is retrieved from the domain registrar InterNIC and includes the length of time the domain has been in existence. Because longevity alone is not a good indicator of an Internet business' reliability, Alexa collects votes from its users on whether they liked a site. Unfortunately, you can't learn why visitors voted as they did. But in one case, for example, a particular travel service developed a large number of "dislike" votes — perhaps a hint that you should consider a different service for your travel needs.

One of the most remarkable features of Alexa is that it can retrieve old Web pages that have moved or changed. When you see the message "404 Not Found" in your browser, you click the Archive of the Web button and Alexa will attempt to display the most recent archived version of that Web page.

This feature grew out of the Internet Archive project, an effort to preserve Web pages for posterity. Officials at Alexa's parent company, San Francisco-based Alexa Internet, say it holds three different "snapshots" of more than 500,000 Web sites that have been collected since early 1996. This information occupies 8 terabytes (8,000GB) of storage — equivalent to all of the data in all of the tapes in a typical video store.

Alexa uses the links between sites, as well as its analysis of how users jump from site to site, to determine other sites that may be relevant to the site you're currently viewing. Clicking Alexa's Where to Go Next button lists these suggestions. The links aren't always pertinent, but they're a start. Alexa supports itself with small advertisements in this box.

You don't even need Alexa to use some of its data. For example, set your browser to <http://widener.alex.com/sitedata/yahoo.com> to see contact and rating information on the Yahoo search engine site. Change yahoo.com to any site you wish.

But the download is well worth it. Go to <http://www.alex.com/download>. I would like to thank Danny Sullivan at <http://search.engine.com> for this suggestion.

Brian Livingston is the co-author of several best-selling Windows books, including the most recent Windows 95 Secrets (IDG Books). Send tips to brian_livingston@intoworld.com. He regrets that he cannot answer individual questions.



TECHNOLOGY

✕ Does Yahoo Still Yahoo?

by Chris Oakes

3.43pm 11.Feb.98 PST

Yahoo may soon have to admit that the very service that made it the Web's Number One site is inherently flawed. As Web growth explodes, critics charge, up to a third of the sites seeking a listing in the popular directory don't get in.

Even Yahoo has long, if quietly, admitted that some sites may take months, even years, to get listed at all. And the impact can be dramatic, as a listing in Yahoo can mean feast or famine when it comes to attracting browsers to a Web site.

Just as it reaches new levels of popularity and business success - the Web's foremost directory service may soon be unable to call itself anywhere near up-to-date.

"They're getting out of the [directory] business, and they should be honest about that," said Louis Rosenfeld, a man who has had personal experience with the difficulty of getting a site listed in Yahoo.

In fact, unexpected reinforcement of his charge can be found in Yahoo's most recent published annual report of 1996, which, in outlining the threats to its business, reads "the Company has from time to time experienced significant delays in the processing of submissions, and further delays could have a material adverse effect on the Company's goodwill among Web users and content providers, and on the Company's business."

"Why is this not a problem anymore?" asks Danny Sullivan, editor of [Search Engine Watch](#), which keeps an eye on the doings of Web directories and search engines.

Submitter Rosenfeld chronicled ongoing problems with his Yahoo listings. While he did get his site listed years ago, his efforts to make corrections when his listings became outdated or incorrect were exercises in frustration.

"I've had terrible problems with the same URLs going back for three years," he said. After waging a "campaign to reach a human" at Yahoo, he found his contact very friendly and responsive - but the fixes they made went only halfway. Today, half his listings remain incorrect, and after his initial go-round, he sees their complete correction as all but hopeless.

Something of a Yahoo competitor, Rosenfeld in fact runs his own service aimed at specialized, rather than generalized, cataloging of topical sites

**TECHNOLOGY
Today's Headlines**

[Intel: Graphics Chip to Make PCs Life-like](#)

[Does Yahoo Still Yahoo?](#)

[QuickTime Blessed by ISO for Streaming Media](#)

[The State of the Net: Crowded, Fragile](#)

[Report: Space Tech NASA Should Develop](#)

[W3C Gives XML Room to Run](#)

[Rockwell Ships Standardized 56-Kbps Chipsets](#)

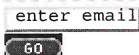
[A Fully Automated Gassing Up](#)

[Think Different: Hack a New Mac OS](#)

[Butterflies Show Path to Cooler Chips](#)

[Sony Prez Calls for a Pan-Industry Tech Standard](#)

[Tracking Down a Mass Killer](#)





BY JERRY TELFER/THE CHRONICLE

In their office at S.F.'s Presidio, Alexa co-founders Brewster Kahle (left) and Bruce Gilliat stood amid their machinery.

Archiving the Internet

Brewster Kahle
makes digital
snapshots of Web

By Carolyn Said
Chronicle Technology Editor

Brewster Kahle is creating the Internet equivalent of the Library of Congress.

The 37-year-old programmer and entrepreneur has been capturing and archiving every public Web page since 1996.

His nonprofit Internet Archive serves as a historical record of cyberspace. His for-profit company, Alexa Internet, uses the archive as part of an innovative search tool that lets users call up "out-of-print" Web pages.

2 Months to Capture

From a 100-year-old, red-roofed office in the Presidio, Alexa's 32 employees send out computer programs that crawl the Internet to find and download Web pages. It takes about two months to capture the entire Web — currently some 300 million pages.

Along with the actual pages, the programs retrieve and store "metadata" — information about each site, such as how many people visited it, where on the Web they went next and what other pages are linked to it.

The Web pages are stored digitally on a "jukebox" tape drive the size of two soda machines. It contains 10 terabytes of data — as much information as one-half the entire Library of Congress.

Like that institution, the Internet Archive doesn't exclude information because it's trivial, dull or just plain weird.

A Virtual Library

"Of course, we've got more pictures of Cindy Crawford than the Library of Congress does," said Kahle. But to create an accurate portrayal of our life and times, it's necessary "to capture all the dreck you could ever want."

Having created a virtual library, the next step was to make a better card catalog. So Kahle and partner Bruce Gilliat started Alexa, named after the ancient Library of Alexandria.

Alexa's search engine uses the Ar-

chive's metadata to help users find information based on the trails of other Internet surfers.

The search engine, available for free at www.alexa.com, is a toolbar that sits along the bottom of a Web browser. It looks at the site a user is currently viewing and suggests other pages by analyzing where previous visitors to that site went next.

Old Sites to View

What separates Alexa from other search engines is that it lets users view sites that have been removed from the Web.

When they encounter the message "404 Document Not Found" users can click on the Alexa toolbar to fetch the out of print Web page from the Internet Archive.

Alexa is supported by advertising, but even the ads relate to users' inter-

"Of course, we've got more pictures of Cindy Crawford than the Library of Congress does."

— BREWSTER KAHLE, Alexa co-founder

ests. A visitor to the Amazon.com Web site might see a Barnes & Noble ad.

"Clearly we need better tools for exploring the Web," says Peter Lyman, head librarian for the University of California at Berkeley and an Internet Archive board member. "Alexa is trying to help us find our way out of the forest by looking for trails where previous people have gone. It's the most promising idea about how we'll search the Internet in the future."

Grander Plans

Available since September, Alexa already has 100,000 users but Kahle has grander plans for it.

"Our goal is to make this part of the infrastructure of the Internet," he said.

One surefire way to achieve that status would be to sell Alexa to a browser company, a search engine company or a major Internet service provider — any of which might be a possibility, Kahle said.

Browser and search firms are snapping up technology that improves Web navigation. Search company Lycos last week spent \$39.75 million for WiseWire, which automatically organizes Internet content into directories and categories. Last month Microsoft shelled out a reported \$40 million for Firefly, which recommends content to Web surfers based on profiles they submit.

Kahle already has a track record of creating next-step Internet technology. In the early 1990s, he developed the Wide Area Information Server (WAIS), the first system for publishing quantities of data in a searchable form on the Internet.

Impressive Background

The New York Times, Wall Street Journal and Encyclopaedia Britannica were among its customers. Kahle later sold WAIS to America Online for \$15 million in 1995.

Besides an impressive programming background, which includes a degree from the Massachusetts Institute of Technology and a stint designing supercomputers at Thinking Machines Corp., Kahle has an abiding interest in traditional media.

His hobby is letterpress printing. Painstakingly aligning individual lead letters by hand to make cards and documents is a far cry from computer automation, "but that's the charm," he said.

His wife, Mary Austin, is the founder and curator of the San Francisco Center for the Book, which runs programs and classes to encourage "all arts of the visible word."

Type Designer's Legacy

They named their 3½-year-old son Caslon after an 18th century type designer. Their 9-month-old son Logan has a family name.

"When the printing press came about, it fostered thousands of tiny presses all over the globe, allowing people in small towns to publish and distribute information. That's what we're finding here on the Web," he said.

"As we move human knowledge from paper to computers, people are getting access to huge amounts of information more easily. But to help organize the Web we have to track what's on it and what's going on over time."

The Chronicle of Higher Education

Academic Today

INFORMATION TECHNOLOGY

March 6, 1998

In Attempting to Archive the Entire Internet, a Scientist Develops a New Way to Search It

Non-profit group uses 'data mining' in effort to preserve World-Wide Web content

By JEFFREY SELINGO

SAN FRANCISCO

For 110 years, Building 116 served unobtrusively as a general store and as quarters for non-commissioned officers stationed at the Presidio, the U.S. Army base south of the Golden Gate. Now that the base has been decommissioned, Building 116 stands out.

Its red-tile roof and cream-colored shiplap siding have been restored, for one thing. And it's one of the few buildings still in use on the 1,480-acre property, which became a national recreation area in 1995. What makes Building 116 unique, though, is what's inside -- a virtual preservation project that aims to create a complete archive of the Internet.

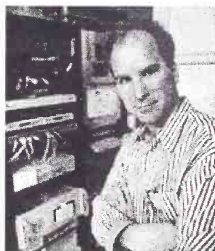
In a back room, a buzz comes from giant computers that are "data mining" the World-Wide Web and Usenet discussion groups, finding and recording pages of the computer network on a nearby digital-tape machine. The computers take complete snapshots of the Web every two months, allowing users to find pages long after their owners have taken them down and let their hyperlinks lapse. So far, the archive has compiled eight terabytes of data -- the equivalent of 800,000 books -- and has recorded at least three snapshots of more than 500,000 Web sites.

Inside those virtual books is the patchwork history of ordinary people: pages of college students long graduated; Web sites of political campaigns since forgotten; early, awkward versions of sites that are now well-known; infamous sites that held our attention for weeks, such as the Heaven's Gate cult's page.

ALSO SEE:

[The World-Wide Web site of the Internet Archive](#), and [a new Internet search engine](#) that can be used to find material in the Internet Archive

[Search The Chronicle](#)[The current Chronicle](#)[Copyright & reprints](#)[Feedback](#)



Fred Mertz for The Chronicle

Brewster Kahle started the non-profit Internet Archive after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million.

Just as the Internet has allowed all kinds of ordinary people to become their own publishers, it has allowed a computer scientist named Brewster Kahle (*left*) to create the non-profit Internet Archive.

Mr. Kahle decided to save the Internet's contents for posterity after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million. The system, which he invented, makes it easier to search electronic data bases.

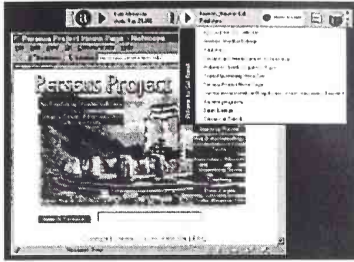
Why an archive? "We need to preserve this heritage," says Mr. Kahle, an affable and enthusiastic 37-year-old who is a graduate of the Massachusetts Institute of Technology. "Or one day, digital anthropologists will wonder if we ever learned anything from the history of other inventions. Remember, nobody recorded television in the early days."

Once he started collecting all that information, Mr. Kahle says, he realized how difficult finding things on line was becoming, with the number of Web sites doubling every six months even as other material falls into neglect. So he set about creating a Web search engine using the technology he developed to manage the massive amounts of data he was collecting for his quirky history project.

The result is Alexa, a search engine operated by Alexa Internet, the for-profit company that is part of the Internet Archive. "This will change the way that researchers use the Internet," Mr. Kahle says.

Alexa is software that can be retrieved free from the company's Web site (<http://www.alexa.com>) and added to a Web browser. Unlike other search engines, such as Yahoo! and Excite, it doesn't rely on word searches. Instead, it watches where its users go on the Internet, and then records that information in a central data base. Based on that information, Alexa can tell a user the most popular paths that other Alexa users have taken from the site the user is visiting at a given time.

It also can suggest other sites offering related material. The top 10 sites pop up in a thin, gray bar near the browser (*see below*) and change as the user moves from page to page.



For example, from the "Perseus Project" (<http://www.perseus.tufts.edu>), a site with an extensive collection of ancient Greek texts in translation, Alexa points the user to sites about classicists and Mediterranean archaeologists at the University of Michigan, sites about publishers and journals available electronically, sites about Hellenistic linguistics, and to "Project Gutenberg," an Internet producer of free electronic texts.

Since October, more than 200,000 people have downloaded Alexa. The service, which Mr. Kahle hopes will soon be fully supported by advertising revenue, is not yet turning a profit. But by the end of the year, he expects it to have a million users. Eventually, he hopes to use the profits from Alexa to finance the gathering of data for the archive.

The advantage of Alexa as a search engine is that it "attempts to be an objective source" for people seeking information. Where conventional links are chosen by a page's creator according to what the creator knows and prefers, Alexa also brings other Web users' knowledge and preferences to bear. The sites recommended in a given search sometimes change, depending on the surfing patterns of Alexa users.

The system has its oddities. If users frequently traveled from the "Perseus Project" to, say, *The New York Times*, the newspaper could be added to the top-10 list of an Alexa user looking at the Perseus site, even though the only thing the two sites have in common is their users. In fact, such a situation has already occurred. From the Perseus site, Alexa suggests -- based on other users' habits -- visits to the sites of Franklin and Marshall College and Bates College. Alexa officials say students at the two colleges probably use the Perseus site in their classes.

"It's sometimes random and not always perfect," Mr. Kahle says. "But if researchers use a traditional search engine, they may miss some of the best sites." A search engine such as Excite, using the keywords "Greek texts" to find sites related to the "Perseus Project," turned up 268,057 matches. "With Alexa, you're bound to hit at

least some of the top sites," Mr. Kahle adds.

Still, one needs a traditional search engine or a specific Web address to get started, Mr. Kahle acknowledges. And Alexa, unlike the Alta Vista search engine and others, can suggest linking only to entire Web sites, not to specific pages within them.

"I don't think of it in the same way as a search engine -- it's a supplement," says Bruce Livett, a reader and deputy head of the biochemistry and molecular-biology department at the University of Melbourne, in Australia. "Alexa gives you relevant sites in the general sense, sites that you sometimes miss because other search engines depend on specific keywords you enter."

Dr. Livett, who has been using Alexa since October, surfs the Web to keep up with the research work of colleagues around the world. "It's competitive work, and I need to know what they're doing." Alexa, he says, has alerted him to research sites that did not turn up in searches using Excite and Anzwers, a search engine designed for Web users in Australia and New Zealand.

Part of Alexa's appeal, he says, is access to the Internet Archive. When Alexa users get a dreaded "404 -- file not found" error, they can click on a button on Alexa's tool bar and pull up the missing page from the archive. Using the archive, Mr. Livett found an audio interview he needed that had been removed from a Web site.

Alexa is "an immediate use for the archive," says Mr. Kahle, adding that the archive is the component that he expects will eventually separate his search engine from the rest of the pack. Alexa also offers a direct link to the *Encyclopaedia Britannica* Web site, allowing users to retrieve reference information without leaving the Web page they are viewing at the time.

As more people begin to use Alexa and the archive, however, tricky questions about copyright and privacy have begun to crop up. An e-mail discussion list for Web publishers recently included a heated debate about copyright issues surrounding old newspaper articles that are part of the archive.

The data-mining computers skip Web pages that require passwords, as well as Web sites protected by the Standard for Robot Exclusion, which blocks search engines from copying pages or directories. Still, some Web publishers said in the e-mail discussion that Alexa officials should be asking on-line newspapers and journals if they want to be part of the archive, instead of forcing them to block Alexa from copying pages.

How the archive will be used in the long term is not clear. Mr. Kahle often mentions the early days of television, when programs were broadcast live and

recording technology was primitive. "When it comes to a point where users have a camcorder recording the Net, then the archive won't be worth it," he says.

In Building 116, the archive is stored in a digital-tape library that looks like a vending machine. The tapes currently have the capacity to hold 20 terabytes of data in all, about as much information as is in the Library of Congress. So much content is being added to the Internet that the archive grows by about a terabyte of data each month. The data-mining computers are able to adjust their site visits to concentrate on those that change most frequently. They will come upon a site, however, only if Alexa users have visited it, if anyone else on the Web has linked to it, or if it is listed with a directory service.

While the archive has been able to keep up with textual information on the Internet -- it is complete from October 1996 to the present -- the effort to collect images is running a few months behind.

Through Alexa, the archive receives about 14 requests for old pages every second. Not bad, its founder says, when one considers that only about 9,000 people visit the San Francisco Public Library on an average day. Alexa also helps researchers by listing facts about the site they're visiting: the address of the individual, company, or other organization that owns the server on which the site is located; how many people have visited the site; how frequently the site is updated; how fast its computers are; and how many pages the site contains. Alexa also allows users to vote for their favorite sites and keeps a running total on each site.

Mr. Kahle says Alexa does not keep individual statistics on its users. Although the search engine tracks the paths of users as they jump from site to site, it does not record users' names. "We don't care who you are," he says. "We just care what path you take."

Mr. Kahle dreams that Alexa could become as popular -- and as profitable -- as search engines like Yahoo! and Excite. And the Internet Archive, he says with enthusiasm, could become part of a large research library, although he's not sure how. "I don't think about the details," he says. "That's why we're doing something now that others thought was impossible, or even crazy."

Copyright (c) 1998 by The Chronicle of Higher Education
http://chronicle.com
Date: 03/06/98
Section: Information Technology
Page: A27

Information Technology

A New Archive and Internet Search Engine May Change the Nature of On-Line Research

Non-profit group uses 'data mining' in attempt to preserve World-Wide Web content

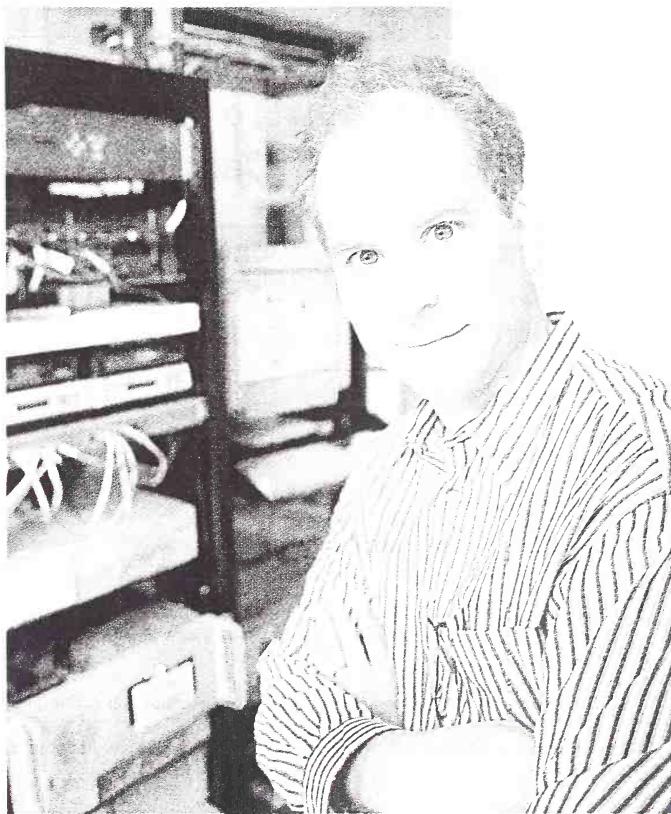
BY JEFFREY SELINGO

SAN FRANCISCO

FOR 110 YEARS, Building 116 served unobtrusively as a general store and as quarters for non-commissioned officers stationed at the Presidio, the U.S. Army base south of the Golden Gate. Now that the base has been decommissioned, Building 116 stands out.

Its red-tile roof and cream-colored ship-lap siding have been restored, for one thing. And it's one of the few buildings still in use on the 1,480-acre property, which became a national recreation area in 1995. What makes Building 116 unique, though, is what's inside—a virtual preservation project that aims to create a complete archive of the Internet.

In a back room, a buzz comes from giant computers that are "data mining" the World-Wide Web and Usenet discussion groups, finding and recording pages of the computer network on a nearby digital-tape machine. The computers take complete snapshots of the Web every two months, allowing users to find pages long after their owners have taken them down and let their hyperlinks lapse. So far, the archive has compiled eight terabytes of data—the equivalent of 800,000 books—and has recorded at least three snapshots of more than 500,000 Web sites.



FRED MURTZ FOR THE CHRONICLE

Brewster Kahle started the non-profit Internet Archive after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million.

DOUBLING EVERY 6 MONTHS

Inside those virtual books is the patchwork history of ordinary people; pages of college students long graduated; Web sites of political campaigns since forgotten; early, awkward versions of sites that are now well-known; infamous sites that held our attention for weeks, such as the Heaven's Gate cult's page.

Just as the Internet has allowed all kinds of ordinary people to become their own publishers, it has allowed a computer scientist named Brewster Kahle to create the non-profit Internet Archive.

Mr. Kahle decided to save the Internet's contents for posterity after selling his previous venture, the Wide Area Information Server, to America Online for \$15-million. The system, which he invented, makes it easier to search electronic data bases.

Why an archive? "We need to preserve this heritage," says Mr. Kahle, an affable and enthusiastic 37-year-old who is a graduate of the Massachusetts Institute of Technology. "Or one day, digital anthropologists will wonder if we ever learned anything from the history of other inventions. Remember, nobody recorded television in the early days."

Once he started collecting all that information, Mr. Kahle says, he realized how difficult finding things on line was becoming, with the number of Web sites doubling every six months even as other material falls into neglect. So he set about creating a Web search engine using the technology he developed to manage the massive amounts of data he was collecting for his quirky history project.

The result is Alexa, a search engine operated by Alexa Internet, the for-profit company that is part of the Internet Archive. "This will change the way that researchers use the Internet," Mr. Kahle says.

THE MOST POPULAR PATHS

Alexa is software that can be retrieved free from the company's Web site (<http://www.alexa.com>) and added to a Web browser. Unlike other search engines, such as Yahoo! and Excite, it doesn't rely on word searches. Instead, it watches where its users go on the Internet, and then records that information in a central data base. Based on that information, Alexa can tell a user the most popular paths that other Alexa users have taken from the site the user is visiting at a given time.

It also can suggest other sites offering related material. The top 10 sites pop up in a thin, gray bar at the bottom of the browser and change as the user moves from page to page.

For example, from the "Perseus Project" (<http://www.perseus.tufts.edu>), a site with an extensive collection of ancient Greek texts in translation, Alexa points the user to sites about classicists and Mediterranean archaeologists at the University of Michigan, sites about publishers and journals available electronically, sites about Hellenistic linguistics, and to "Project Gutenberg," an Internet producer of free electronic texts.

Since October, more than 200,000 people have downloaded Alexa. The service, which Mr. Kahle hopes will soon be fully supported by advertising revenue, is not yet turning a profit. But by the end of the year, he expects it to have a million users. Eventually, he hopes to use the profits

from Alexa to finance the gathering of data for the archive.

'OBJECTIVE SOURCE'

The advantage of Alexa as a search engine is that it "attempts to be an objective source" for people seeking information. Where conventional links are chosen by a page's creator according to what the creator knows and prefers, Alexa also brings other Web users' knowledge and preferences to bear. The sites recommended in a given search sometimes change, depending on the surfing patterns of Alexa users.

The system has its oddities. If users frequently traveled from the "Perseus Project" to, say, *The New York Times*, the newspaper could be added to the top-10 list of an Alexa user looking at the Perseus site, even though the only thing the two sites have in common is their users. In fact, such a situation has already occurred. From the Perseus site, Alexa suggests—based on other users' habits—visits to the sites of Franklin and Marshall College and Bates College.

Alexa officials say students at the two colleges probably use the Perseus site in their classes.

"It's sometimes random and not always perfect," Mr. Kahle says. "But if researchers use a traditional search engine, they may miss some of the best sites." A search engine such as Excite, using the keywords "Greek texts" to find sites related to the "Perseus Project," turned up 268,057 matches. "With Alexa, you're bound to hit at least some of the top sites," Mr. Kahle adds.

Still, one needs a traditional

Circ - (6) 52 734 (\$ 55.804

MARCH 23, 1998



Bacon's

3985-A



The Shareware Report

Alexa can find sites you'd like

We've heard so much about this Alexa tool that we had to try it.

Alexa, according to the help file, is "a free Internet navigation service that helps you surf the Web smarter, faster and easier. The service helps you determine the value of web sites you visit through a variety of statistics and facts about each site and then recommends links of where you might want to go next."

It certainly does that, perhaps ad nauseum. The toolbar, which Alexa places at the screen's bottom, has many choices. If you find it intrusive, minimize it on the taskbar, or select autohide. It suggests, according to your current Web site, where you might go next. It even lets you vote on whether you like a Web page, then archives it. You can also add links to the list.

Alexa provides a button to the Encyclopedia Britannica site, handy for quick reference, especially with its search window. If you run both major browsers, it works with either, not just your default.

A better feature is the archival reference: if you receive the dreaded "404 URL not found" response in your browser, ask Alexa to retrieve the last valid page for that address. Alexa is one of the best integrated desktop/Internet navigation accessories we've seen.

Where to find it:
www.alexa.com

Price: Free

System Requirements: A PC running Windows 3.11 or 95 and an Internet

Bill Dubie and Dave Sciuto co-host "The Computer Report" on WCAP AM-980 in Lowell on Sundays at 7 am. Check out their Website at <http://thereport.com/> or con-

search engine or a specific Web address to get started, Mr. Kahle acknowledges. And Alexa, unlike the Alta Vista search engine and others, can suggest linking only to entire Web sites, not to specific pages within them.

'A SUPPLEMENT'

"I don't think of it in the same way as a search engine—it's a supplement," says Bruce Livett, a reader and deputy head of the biochemistry and molecular-biology department at the University of Melbourne, in Australia. "Alexa gives you relevant sites in the general sense, sites that you sometimes miss because other search engines depend on specific keywords you enter."

Dr. Livett, who has been using Alexa since October, surfs the Web to keep up with the research work of colleagues around the world. "It's competitive work, and I need to know what they're doing," Alexa, he says, has alerted him to research sites that did not turn up in searches using Excite and Anzwers, a search engine designed for Web users in Australia and New Zealand.

Part of Alexa's appeal, he says, is access to the Internet Archive. When Alexa users get a dreaded "404—file not found" error, they can click on a button on Alexa's tool bar and pull up the missing page from the archive. Using the archive, Mr. Livett found an audio interview he needed that had been removed from a Web site.

Alexa is "an immediate use for the archive," says Mr. Kahle, adding that the archive is the component that he expects will eventually separate his search engine from the rest of the pack. Alexa also offers a direct link to the *Encyclopaedia Britannica* Web site, allowing users to retrieve reference information without leaving the Web page they are viewing at the time.

QUESTIONS OF COPYRIGHT

As more people begin to use Alexa and the archive, however, tricky questions about copyright and privacy have begun to crop up. An e-mail discussion list for Web publishers recently included a heated debate about copyright issues surrounding old newspaper articles that are part of the archive.

The data-mining computers skip Web pages that require passwords, as well as Web sites protected by the Standard for Robot Exclusion, which blocks search engines from copying pages or directories. Still, some Web publishers said in the e-mail discussion that Alexa officials should be asking on-line newspapers and journals if they want to be part of the archive, instead of forcing them to block Alexa from copying pages.

How the archive will be used in the long term is not clear. Mr. Kahle often mentions the early days of television, when programs were broadcast live and recording technology was primitive. "When it comes to a point where users have a camcorder recording the

Net, then the archive won't be worth it," he says.

In Building 116, the archive is stored in a digital-tape library that looks like a vending machine. The tapes currently have the capacity to hold 20 terabytes of data in all, about as much information as is in the Library of Congress. So much content is being added to the Internet that the archive grows by about a terabyte of data each month. The data-mining computers are able to adjust their site visits to concentrate on those that change most frequently. They will come upon a site, however, only if Alexa users have visited it, if anyone else on the Web has linked to it, or if it is listed with a directory service.

14 REQUESTS PER SECOND

While the archive has been able to keep up with textual information on the Internet—it is complete from October 1996 to the present—the effort to collect images is running a few months behind.

Through Alexa, the archive receives about 14 requests for old pages every second. Not bad, its

**"If researchers
use a traditional
search engine,
they may miss
some of the best
sites."**

founder says, when one considers that only about 9,000 people visit the San Francisco Public Library on an average day. Alexa also helps researchers by listing facts about the site they're visiting: the address of the individual, company, or other organization that owns the server on which the site is located; how many people have visited the site; how frequently the site is updated; how fast its computers are; and how many pages the site contains. Alexa also allows users to vote for their favorite sites and keeps a running total on each site.

Mr. Kahle says Alexa does not keep individual statistics on its users. Although the search engine tracks the paths of users as they jump from site to site, it does not record users' names. "We don't care who you are," he says. "We just care what path you take."

Mr. Kahle dreams that Alexa could become as popular—and as profitable—as search engines like Yahoo! and Excite. And the Internet Archive, he says with enthusiasm, could become part of a large research library, although he's not sure how. "I don't think about the details," he says. "That's why we're doing something now that others thought was impossible, or even crazy." ■

* Links to the Internet Archive site can be found on *The Chronicle's* World-Wide Web site at:

<http://chronicle.com>

The Digital Attic: An Archive of Everything

JAMES GLUECK
FAST FORWARD

Before the Internet, you lost data nearly every time you upgraded your computer. Now you couldn't get rid of that embarrassing E-mail if you tried.

YOU PROBABLY HAVEN'T SPENT MUCH time worrying about what will happen to your Web site when you're dead. That's all right. David Blatner is worrying for you. "I keep thinking," he says, "if my grandparents had built a Web site, wouldn't I want it archived and available on the Net in the years to come for their grandchildren?" So he is ready to help with his new Web-preservation organization in Seattle: Afterlife.

Meanwhile, in central Ohio, a site called Orphans of the Net is "rescuing" some Web pages that have been abandoned or shut down — generally shrines for minor celebrities. If you're looking for old publicity photos of Kimberly Williams or Renée Zellweger, rest assured that they have been preserved on line.

These modest salvage jobs notwithstanding, many of the world's librarians, archivists and Internet experts are warning that the record of our blooming digital culture is heading for oblivion, and fast. They note that we have already begun losing scientific data and business records — stored on ancient punch cards or written in dead computer languages or encoded on decaying Univac Type II-A magnetic tape. (Just try to find a Univac tape reader when you need one.)

In the electronic era, we are stockpiling our heritage on millions of floppy disks, hard drives and CD-ROM's. These flaky objects go obsolete dismayingly fast, with new technologies rolling in on product cycles as short as two to five years.

"There has never been a time of such drastic and irretrievable information loss," says Stewart Brand, creator of the "Whole Earth Catalog" a generation ago and an organizer of a sobering conference earlier this year called "Time and Bits." Our collective memory is already beginning to fade away, many of the participants believe. Future archeologists will find our pottery but not our E-mail. "We've turned into a total amnesiac," Brand says. "We do short-term memory, period."

The information-storage medium of the past couple of millennia — for words not writ in

stone, anyway — has of course been paper. Paper does decay with time, and it is fragile. One big fire at the library at Alexandria in 391 A.D. destroyed a calamitous piece of the ancient world's heritage. But to some people, paper starts to look good.

"Paper at least degrades gracefully," says Brand nostalgically. "Digital files are utterly brittle; they're complexly immersed in a temporary collusion of a certain version of a certain application running on a certain version of a certain operating system in a certain generation of a certain box, and kept on a certain passing medium such as 5¼-inch floppy." If a company has digital business records a mere decade old, what are the chances that it has also stored a vintage 1988 personal computer, DOS 2.1, and the correct version of Lotus 1-2-3?

Some companies have begun "refreshing" their aging records by continually copying them onto new storage media, using new software. Refreshing isn't easy, and most institutions have not yet realized that it may be necessary. Whatever media they used to save their digital information, they will not be able to read it without a machine — a finicky antique, most likely. With paper, all you need is your eyes.

Perhaps the speed and richness of the Internet have lulled us, letting children in Boise read Census data in Washington and oral history in Hiroshima. Words swim instantly across the network, not caring about the mileage, and we don't exactly feel information-deprived. But are we sacrificing longevity to gain glut?

"Back when information was hard to copy, people valued the copies and took care of them," says Danny Hillis, co-founder of Thinking Machines Corporation and now vice president of research at Disney. "Now copies are so common as to be considered worthless, and very little attention is given to preserving them."

It's scary. And yet . . .

Anyone wandering through the Internet might begin to feel that memory loss isn't the problem. Archivists are everywhere, in fact — official and self-made. On Sunday, July 3, 1994, I

played a hand of bridge that would be best forgotten — but no, the leading on-line bridge service, OKBridge, has recorded every detail of the bidding and card play in each of the seven million hands played since the beginning of that year.

Likewise, any silly message that you broadcast to any Usenet newsgroup is now being stored, for eternity or some approximation thereof, by a variety of commercial services. No matter that you gave your last posting a mere five seconds' thought; you should be prepared to hear your biographer read it back to you in your dotage.

Most people, unfortunately, don't have posterity in mind when they fire off their little notes. Internet communication seems so spontaneous and personal. Will people really want future employers to dig up all the messages they've been posting to alt.dead.porn.stars and soc.support.depression.manic? Sometimes, as the years go by, privacy demands a gentle forgetfulness.

Many people sitting at company workstations toss off E-mail as casually as they speak — gossipy E-mail, secretive E-mail, snide E-mail, raunchy E-mail, E-mail meant to self-destruct after serving its instant purpose. But it lives on, as corporate lawyers and prosecutors have realized. Neither sender nor recipient can delete it reliably. To the lawyers' occasional horror — here comes the subpoena! — it lingers on disk drives and backup tapes like a late-night guest who has forgotten how to leave.

The biggest proprietor of archivable data is the Federal Government, struggling to preserve the records it generates daily on an uncountable scale. It is a matter of current litigation whether every piece of governmental E-mail must be preserved as a "Federal record." Either way, the task of the National Archives and Records Administration is

monumental. "What we're looking at is growth that there's no way we can deal with, using any known technique or resources we can get," says Ken Thibodeau, director of the Archives' electronic records programs.

"Digital information technology is creating major and serious challenges for how we're going to preserve anything of our culture and our history," Thibodeau says. "It's also creating opportunities: we'll be able to preserve and use a lot more information than ever before." Pity the poor historian, though. The Clinton White House's E-mail alone figures to be 8 million files.

Meanwhile, in its unofficial way, the Internet is transforming the way information is stored. The traditional function of libraries, gathering books for permanent storage or one-at-a-time lending, has been thoroughly confused. Archiving of the on-line world is not centralized. The network distributes memory. There is a kind of self-replication at work, with data employing humans in the effort to spread and reproduce.

Web site by Web site, the data seem as frail as skywriting — smoke in the breeze. Brewster Kahle, inventor of some of the best Internet search systems, estimates the average lifetime of a Web page at 75 days. He has created the Internet Archive, though, to store periodic snapshots of almost the entire World Wide Web. It maintains pages lost or shut down by their owners. It amounts to about eight terabytes of data. ("Tera-" is 1,000,000,000,000. Get used to it.)

Brand and his fellow Casandras have a point, and they are focusing attention on some new, practical issues. Who, if anyone, will decide which parts of our culture are worth preserving

for the hypothetical archaeologists of the future? Can any identification scheme help readers distinguish true copies from false copies in the on-line world's hall of mirrors? What arrays of optical or magnetic disks might provide reliability and redundancy for more than a few years of storage? Still, hope comes from the simple truth that the essence of information does not lie in any technology, new or old. It's just bits, after all.

In the world before cyberspace, countless bridge hands were played and words spoken and the memory vanished like vapor into the air. Think of all that data, dissolving no sooner than it was formed. Once in a while people managed to snatch a bit back from the ether, with pen on paper or, later, audio- and videotape. They succeeded in saving a fair portion of what was worth saving: the speeches of Lincoln (the major ones), the poetry of Shakespeare (but not quite reliably), the plays of Sophocles (except the lost ones) and a few dozen terabytes more.

Everything is different now. The Internet turns much of humanity into a sort of giant organism — an intermittently connected information-gathering creature — and really, amnesia doesn't seem to be its fatal flaw. This new being just can't throw anything away. It is obsessive. It has forgotten that some baggage is better left behind. Homo sapiens has become a pack rat.

Shed tears if you must for the backup tapes already demagnetized. You'll have many opportunities. Just last October, the Daioh Temple of Rinzaï Zen Buddhism held a "memorial service for lost information" in Kyoto and on line. Of course, the details are lovingly preserved, in English and Japanese, at its Web site. ■

Circ - (M) 7,500
8a

MARCH 14, 1998

**Bacon's**

3985-A

Brewster Kahle downloaded an entire library for you

If you have a lot of time on your hands, and about 8 terabytes (8000 gigabytes) of hard drive space, you could download the entire World Wide Web. That's equivalent to all the books in the Harold Washington library.

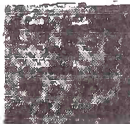
You'd probably be bored with the whole idea by the time you finished—assuming you were still alive. Fortunately, Brewster Kahle has already done the job for you.

Kahle was an early Internet developer, creator of Wide Area Information Server (WAIS), a system that preceded HTML (hypertext markup language) as a way of publishing and indexing Internet information.

From a project to preserve Web pages for posterity—the Internet Archive—Kahle's company, Alexa Internet, developed a new kind of Internet tool that may not change the way we use the Web, but will at least add a powerful new dimension to it.

Named after the lost library of Alexandria, Alexa is a browser add-on that runs separately, and consists of several components. Installing Alexa in Windows 95 or NT will place its toolbar in the tray at the bottom of your screen when Windows is booted.

When you start your browser the toolbar shows the site name. Click it to get a variety of interesting information about the site, including the name and address of its registered owner, and the date



COMPUTING

BY JERRY MAIZELL

on which it was registered.

You can also "vote" on whether you like or dislike the site, and there's a note of how many other Alexa users voted one way or the other.

This is useful if you have some question about the site's integrity, such as when you wish to purchase something from it.

Alexa also provides ratings on the site's popularity, citations of reviews if available, its typical loading speed, a "freshness" rating (how often it's updated), and how many links and pages it has.

Alexa also tries to recommend other sites similar to the one you're viewing, by providing links to sites that were followed by other recent visitors.

Because Alexa Internet has downloaded (and occasionally updates its "snapshots" of) the entire Web, its archive icon can bring up the most recently archived version of many sites that have moved, or no longer exist.

This can relieve the frequent frustration of encountering the dreaded "404" (page not found) error message.

There is a handy link to the Encyclopedia Britannica, and the Mernam-Webster online

dictionary and thesaurus.

Supported by small, non-intrusive ads that appear on its toolbar, Alexa's price is right free. It's a fast 1.2mb download at www.alexa.com.

You can test its information service without downloading it: set your browser to <http://widener.alexa.com/sitedata/whitehouse.gov> for a sample (substitute any site you wish instead of whitehouse.gov).

With Windows 95 be sure you have a version of the file `Kernel32.dll` dated later than Feb. 2, 1996. Use the `Find` command on your Start menu to locate it and check its date. An updated version can be downloaded from Alexa's site.

Alexa works with Netscape or Internet Explorer version 3 or later. It also works with my new favorite browser Opera (NNN Feb. 28).

Serious surfers won't browse without it.

* * *

To: Cynthia Lohr - Here is the article about Alexa! - Carl

SUNDAY, MARCH 1

on-line

Alexa is your window to the whole wide web

Get maximum mileage out of search engines

By Carl Lum
Contributing writer

YOU'VE FINALLY SET UP YOUR COMPUTER to access the Internet, but how do you find anything in this vast virtual "library?" Search engines are the logical places to start. Yahoo! (www.yahoo.com) is a household name in Internet circles, as are Excite (www.excite.com) and Alta Vista (www.altavista.com).

Sometimes these engines are a little too thorough. A typical search may result in thousands of hits. It takes a determined effort to separate the wheat from the chaff. Browsing through categories offered by many search services may help, but it's tedious and only helpful if you're looking for something general.

This is where Alexa comes in. In an industry where "cool" and "unique" are twin Holy Grails, Alexa is an ambitious undertaking that's free and designed to put human intelligence into Web navigation. Like any good guide, it eagerly informs the user about the site visited and quickly rattles off a list of similar sites.

How does Alexa do this? This is where the ambitious part comes in: Alexa takes a "snapshot" of the entire Web every 30 to 60 days, recording the entire World Wide Web. It takes about two TB (terabytes) of disk space to store the Web. Alexa has about eight terabytes in its archives. For comparison, a typical public library contains about three TB of information, a video rental store holds about eight TB, while the entire computer generated *Toy Story* movie takes up half a terabyte. (A terabyte is 1,000 gigabytes, or 1,000,000 megabytes.)

In addition to this huge archive, Alexa uses anonymous "user paths." Every time an Alexa client is used, it records where the user went, how long was spent at each site,

and if the user gave the site a thumbs up or thumbs down. It also notes how quickly the site loads, how often it's updated, and who owns the site.

Alexa uses this information and other sources to generate its list of "Where to Go Next" sites. And unlike the thousands of hits returned by a search engine, Alexa's list is usually no more than a dozen highly relevant items.

To use Alexa, download and install the client software from www.alexa.com. During installation, provide general demographic information about yourself—Alexa supports itself solely from advertising revenue. The option "Rather not say" is thoughtfully provided if you prefer not to.

When you begin surfing, Alexa springs into action, displaying useful information in a toolbar (or "dashboard") below your web browser.

- **Where You Are** displays the web site's popularity, who owns the site, and other nifty information;
- **Where to Go Next** recommends similar sites to look at;
- **Archive of the Web** retrieves missing web pages.

• **Desktop Reference** provides instant access to the *Encyclopaedia Britannica* and the *Merriam-Webster Dictionary and Thesaurus*.

I found Alexa responsive and solid for an early version software. Its site recommendations were interesting: While at Microsoft's site, Alexa suggested that I visit Apple, Sun, and Netscape—three of the most rabid anti-Microsoft companies out there. However, there was no mention of visiting Microsoft when I went to sites for Apple and Netscape.

Alexa requires a 486/33 running Windows 95/NT. A Macintosh version is in development.

carllum@uniserve.com

Spring Break Theatre

Performs Evenings

"On the Waterfront"

"For Art's Sake"

"Sudden Death"

"The Beautiful Game"

"Theatre Square"

The Waterfront

or Ticket

\$7.50 per person (all ages)
Groups: (10 or more)

Arts
Theatre Co.

CAROLINE THEATRE
Company of the Arts
at Glenview High School

Four professional companies with international
of theatre for young audiences right

courier

LYNX

Glenview High School

YEAR-END

Put Your Office
Productivity On
The Fast Track
With Minolta

GILL & CO.



New products help coordinate several search engines and yield more efficient results.

By JOHN HOKKANEN
SENIOR ATTORNEY AT THE NATIONAL LAW JOURNAL

WITH THE EXPANDING Internet and the increased use of electronic data storage in the modern law office, the goal of maximizing the knowledge relevant to law practice has become increasingly difficult. Fortunately, as the Internet and Intranets have become platforms for publication and collaboration, a host of new, mostly new search tools have appeared.

Past discussions of search technologies have focused on which major search engine was the best or most comprehensive. Now new, small desktop search tools seek to leverage the big search engines' repositories and tackle different kinds of problems. The new search widgets seek to make searching multiple Internet/intranet search engines a less painful task. Searching numerous engines with different interfaces can be time-consuming. The lists resulting from the various searches frequently include many of the same documents, and users must identify and scan duplicate hits in order to move on to uninspected documents. In addition, there is often no way to store one's search query.

An additional problem relates to a group's ability to create a customized knowledge base or data collection that may be shared and easily searched. Full-text retrieval may be a great way to move around in undifferentiated data, but a group with a particular domain expertise, such as environmental law, might wish to collect the various Internet and intranet resources and apply structure to such content. A collaborative method of creating, maintaining and structuring a knowledge base with a friendly folder would be useful. Such knowledge bases could also be cloned and placed on a secure extranet for a client, and this additional service could be offered at no cost. Having collectively populated a knowledge base with materials, a group might also benefit from having users provide additional feedback. For example, a user might wish to provide ratings of particular documents to reflect their overall value or annotations to identify specific items of interest.

To the extent that usage of the collection is monitored, the system might also assess which documents may be of more or less value to users looking for particular items, and the evaluations or rankings can provide a foundation for group filtering algorithms. All of these ideas, have already been implemented to varying degrees in existing Internet technologies. Reviewing some of these products may provide one with ideas about which features could be useful for an office's next legal practice system.

Applications That Interact

In creating a practice system, a law office may wish to combine some of these technologies, using both off-the-shelf (the) and custom pieces. One common feature of these new search tools is the use of a separate window or application to create a persistent search interface. The applications come in two flavors, more or less. The first is a separate Windows application that can interact with the Web browser. The second variety implements a solution completely within the Web browser with a separate window.

Examples of the Windows applications category are Quarterdeck's Web Compass,¹ Web Ferret,² SSSpider³ and Alexa.⁴ These applications interact with the major search engines in various

ways, and each offers interesting capabilities. For example, Web Compass Web Ferret and SSSpider will search multiple search engines and combine the resulting sets into a single set with duplicates removed. Web Compass allows the user to add additional search engines, such as Intranet, and schedule automatic checks for new information.

Alexa is an interesting piece of proprietary groupware. It lets the user run the application and use the Web browser. It is not a tool for actively searching major engines; instead, as one browses, the program communicates with the Alexa server. The Alexa toolbar provides different statistics about the Web site and shows related sites. Alexa users can even review a site by giving it a "thumbs up." This passive relating of one site to another opens up completely different avenues and allows the searcher to perform collaborative filtering of sites.

Melting Point is groupware of a different sort, allowing an office to collect rela-

tionship information about documents in PCDocx, in MSExchange or on an Intranet. The information is stored in a server data base, allowing colleagues to share and search the relationship information authored by others.

Some new search engines are providing the capabilities of these applications. For example, Northern Light will provide a folder-based view of the results of a search. Meanwhile, Inference Find⁵ searches six of the major search engines, removes duplicates and provides a relevancy-ranked list.

Browser-Based Tools

A Navigator/Explorer program similar to the Windows applications is WebTurbo⁶ which is distributed at no cost. This tool is similar to the Windows applications in that it searches major search engines and removes duplicates from the resulting list; documents on the list may be previewed. The application is free because the company sells advertising in

the application window.

Browser-based tools do not stop with centralized server resources. For example, at IBM's Web site⁷ when one runs a keyword search, a separate search tool window is spun off to help the user navigate the site. This is similar to a tool known as the Search Satellite,⁸ a Java script browser tool that gives the user a consistent interface for searching the major search engines. Alternatively, Don T's search tool⁹ provides a separate window, though this tool does not integrate the various search engines into a single interface, as does the Search Satellite.

The InfoMinder¹⁰ a JavaScript/Java-based tool similar to the Search Satellite interacts with a centralized server data base. In addition to being able to categorize search engines in a folder metaphor and then submit queries as the Search Satellite does, InfoMinder can also store URLs and full-text index the document for subsequent searches. With this tool, a

[SEE WIDGETS PAGE B12]

National Law Journal (N.Y.N.Y.)
April 13, 1998

[WIDGES FROM PAGE 81]

group can collectively assemble a knowledge base of Internet/intranet documents. As with Alexa, users can provide rankings of documents, and this information is included on the lists of results.

At the same time, developments have also occurred in the area of collaborative filtering.⁽¹⁾ This technology demonstrates how a very small set of input data, placed in the context of a large number of similar users, can generate some powerful collaborative filtering results.

One can see this technology in action at the MovieCritic⁽²⁾ site where, after the user rates a dozen movies, the engine provides recommendations based upon the ratings of people who liked and disliked the same movies. Because this approach can provide meaningful results with a small input set, it shows that human preferences can cluster and that clustering can subsequently be used. For those with doubts about where technology is headed, this site provides an interesting vision of the upper limits of more or less passive, customer-driven tools.

Collaborative filtering mechanisms may play an important role in leveraging lawyer work product as now associates look for good work product created by others. There is some risk that the guidance that a collaborative filtering system provides might go beyond passive suggestion and herd a new user toward documents that are frequently used, overlooking a more obscure but more penetrating document. But full-text searching of the collection by the user may provide a check on such automation.

Finally, one site offers a way to create a feedback mechanism to one's e-mail account. The Informant⁽³⁾ allows users to specify Web sites, and the server then e-mails users with updated information.

The convergence of search-engine, groupware and collaborative filtering technologies within the Web environment demonstrates the possibilities for development of new legal practice systems. Full-text, relevancy-ranked results may be complemented by authored, hierarchical, context-based mechanisms of categorizing information.

By decentralizing the authoring of this context data, the applications become interactive groupware, allowing an entire group to share both the work of developing such a resource and its benefits. The ability to create systems with feedback mechanisms furthers this groupware approach, and helps create markers that indicate which documents are useful, current or of little consequence.

Collaborative filtering technologies that monitor the use of the system and provide its users with passive suggestions about other documents can make systems even more usable to lawyers with limited computer skills. In the process, users may be directed to documents that they may not have known exist.

When applications or their server-based components are multi-aware, the systems might notify users when new materials are added and which would qualify for a user's search or be of interest when a user is perusing the collection.

(1) <http://www.schaffrquick.com/widges/products/wc200/>

(2) <http://www.moviecritic.com/movierest/>

(3) <http://www.kryptech.com/download.htm>

(4) <http://www.alexa.com>

(5) <http://www.intelligence.com>

(6) <http://www.wchirolog.com>

(7) <http://www.ltm.com>

(8) <http://http://www3.pair.com/journey/search/>

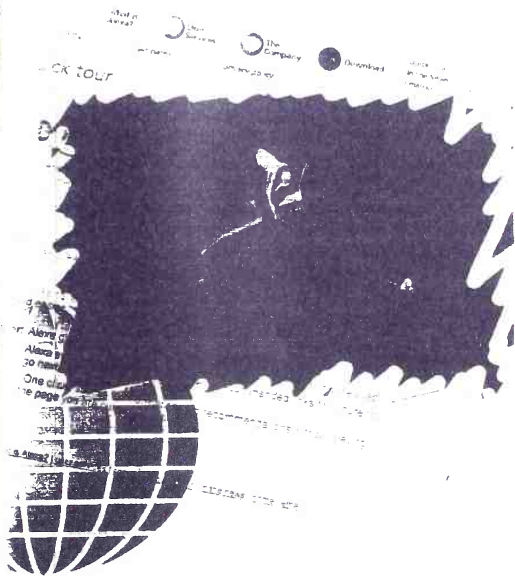
(9) <http://www.genetics.com/ILConValley/My/8030/>

(10) <http://www.lpr.com>

(11) <http://www.ltkonltd.com>

(12) <http://www.moviecritic.com>

(13) <http://informant.darnmouth.edu/>



"Fun, Funded and Fearless"

Alexa Internet Wants to Change the Future
and Make Your Life Easier

By Tracy Levine

Alexa Internet is a unique firm that has created a very important tool you can download today—for free. The company has its offices within the Presidio, in a historic San Francisco building dating back to circa 1885. The exterior is unassuming, but inside—the excitement and energy levels are high.

Led by Brewster Kahle, the corporate culture manifests innovation, ideals, vision and an ethical consciousness. Like the executives of many companies who have chosen the Presidio as their headquarters, the Alexa management team is not purely profit-driven. Rather, they are progressive thinkers who want to be a part of a better tomorrow.

"Information should be shared—not the plaything of billionaires or lost completely," explained Alexa president Brewster Kahle. "We are creating open, civic spaces and fostering the sharing and archiving of knowledge."

Kahle claims that Alexa Internet is the first company to offer Web navigation that actually learns from its users. (More about this point later!) Founded in April 1996 by Brewster Kahle and Bruce Gilliat, the company provides a solution to a major problem of the Internet: the inability to find information quickly and easily.

Alexa is a free service that helps you surf the Web faster and much more efficiently. It works in conjunction with Web browsers and is displayed as a toolbar on your desktop.

**"We are creating open, civic
spaces and fostering the sharing
and archiving of knowledge."**

Taming the Nightmare

As the Internet has grown to overwhelming and unmanageable proportions, navigating has become a total nightmare. Since search engines—even the good ones—rely on keywords to locate information, you can be thrown, unwittingly (and unwillingly!) into a melting pot of information that is completely irrelevant to your search. Wading in this quagmire of meaningless data is time-consuming and frustrating. Alexa cuts through all that and effectively helps you target your search and even suggests links related to your personal quest.

"We operate purely from the user's perspective and are focused on helping people make sense of the Web. We think the Web is worth it," said Kahle.

Unfortunately, as the Web has become more convoluted, the game has become more complex while the stakes (and potential profits) have increased as well. Cynthia Lohr, public relations manager for Alexa, insisted: "Search engines are evolving into families of online media properties. Their primary purpose is no longer to help you find information. They've become malls—trying to keep you in their stores as long as possible."

Alexa Is Helpful

Alexa is certainly very helpful. I used it to research some stories and found the time-saving aspect to be most valuable.

The key features of Alexa are:

- **Where you are**—gives you worthwhile information, such as who the site is registered to, how frequently it is updated, plus ratings and reviews from publishers and other third-party sources.
- **Where to go next**—delivers suggested sites to further flesh out your search.
- **Desktop reference services**—integrates, into your browser, *Britannica Online*, including the new *Britannica Internet Guide*, as well as the *Merriam-Webster Online* dictionary and thesaurus.
- **Archive of the web**—lets you access Alexa's extensive archive of the Web to reach previously (annoyingly) unavailable "404-Not found" pages.

This last feature is most intriguing. I, for one, did not realize that information is simply vanishing all the time. It's not something we really spend a lot of time thinking about. Isn't someone in charge of keeping all that information? Well...no. This fact was an impetus for the founding of Alexa.

How It All Started

Brewster Kahle was a key pioneer in the Internet publishing industry. In 1989, he invented Wide Area Information Servers (WAIS), an early Internet publishing system that worked with Dow Jones, the *New York Times* and *Encyclopedia Britannica*. He sold WAIS to America Online in 1995 for \$15 million.

How did Kahle segue into his current enterprise? "It was exciting to play a role in the Internet publishing movement. Building an information community seemed to be the next logical step. We're building something really neat here that will, we hope, make the world a little better for it," he said.

Kahle and Gilliat have embarked on a very ambitious mission. They are archiving the Web (yes, the whole Web) in a database for...essentially, mankind. Alexa began "collecting" the Web in early 1998, and it now has a minimum of four snapshots of over 800,000 Web sites.

So how vast is Alexa's archive? It has already surpassed Alexandria in size and is now equal to the Harvard University archives. Company officials say they expect it to pass even the Library of Congress in the coming months.

Well, it's certainly comforting to know that someone has taken a leadership role in protecting and preserving all that information—not only for us but also for our future generations. Yet, Alexa wants to be more than just a massive library. Its founders have raised the bar to an intriguing level. Information is being shared. Information is being saved. And, it's getting richer as more users work with it.



About That Learning Process

Alexa actually learns (through an intricate form of data analysis and capturing usage paths) from the people who use it. So, it follows that as the number of users increases, so does the value of the information. Kind of like a good chain letter!

Cynthia Lohr described it this way: "When you walk on a path through the woods, you are benefiting from the explorations of those who went before you—finding the best way up the mountain, or down to the sea. Alexa seeks to do the same for the Web. You gain the benefit of everyone's experience." For example, if you were looking for background on Hong Kong, your initial personal search might lead you from the official tourist site to the discovery of sites for the daily English-language newspapers published in that Chinese city. The Alexa software would note the steps in your search and suggest those newspaper sites to later Alexa users inquiring about Hong Kong.

"We operate purely from the user's perspective and are focused on helping people make sense of the Web. We think the Web is worth it."

Making It Pay

Alexa does profess to have a cause, rather than just a sales goal. But, then again, even those bound by altruism have to eat! There have been 300,000 downloads of Alexa so far—all free of charge. The question is, how are its owners making a profit?

The company was seeded with some of the gains from the founders' WAIS, Inc. transaction. It continues to grow through its unique advertising strategy. Alexa offers advertisers the singular ability to advertise in a very targeted way. The message is accessed through the Alexa toolbar displayed on the user's screen. Context-oriented advertising allows advertisers to target specific Web sites, geographic regions and even the time of day. A golf club vendor, for example, would benefit from advertising on Web pages that reach golfers.

Alexa has 50 advertisers so far, and the marketing department says they've generated a lot of interest from additional organizations that want to do a better job of targeting their Web advertising.

The company has some impressive goals, and because they are in a ground-breaking mode, they have been able to attract a "great group of people" (27 on the team and growing). Success and creative satisfaction are what the Alexa folks are shooting for.

But, what does Brewster Kahle predict for the future of the Internet?

"Video! It's here now, but soon it will be much more prevalent. It will change everything. Just think of how you'll be able to cut down on business travel. Rather than attending a trade show, you'll be able to preview products through online video. The world of education will be greatly enhanced too and add a whole new spectrum of possibilities!"

Kahle has a personal technological goal too. He is trying to set up an unprecedented "sonic speed" networking system that will unite all companies and residents in the Presidio. After experiencing this executive's drive and enthusiasm, I expect this concept will soon be a reality.

To learn more about Alexa or to test-drive it yourself, you'll find the company at www.alexa.com. For information about the firm or its advertising opportunities, you can call Cynthia Lohr at (415) 561-6786.

Context-oriented advertising allows advertisers to target specific Web sites, geographic regions and even the time of day.

OH-W2170

WEEKLY NEWSPAPER

THE SUN STAR
Strongsville, OH

Circ - (M) 11,712
Th

FEBRUARY 19, 1998



Bacon's

8785--

SOUTHWEST I

Internet just keeps on growing, some cool tools help you surf

By **DENNIS SEEDS**
Editorial Systems Manager

A new freeware program called **Alexa** is making news in the Internet world — a world that counts 56 million Americans as users. Alexa is an add-on to your Internet browser that gives you a rating of the Web site you are viewing, and can take you to lost Web pages that have been erased.

The Alexa company has made a backup copy of the entire World Wide Web. That doesn't sit well with some publishers who have taken their articles off line not intending them to be available any more, and some controversy is brewing.

Alexa also has shortcuts to reference material, such as a dictionary, thesaurus and encyclopedia.

In the ratings area, it will tell you who has registered the domain, how often the site has been visited, how often it is updated and how many pages are on the site.

You can download a copy at www.alexa.com.

Here are this month's reviews of Greater Cleveland Web sites:

Heather Hill Health & Care Center

New to the Web is this site for Heather Hill Hospital. Health & Care Center in Chardon, a non-profit facility for the frail elderly and physically disabled.

The Web site has interactive features, such as moderated discussion forums and a cool virtual reality tour of the facilities. The tours are 360-degree pans around three options: the interior common area of the Corinne Dolan Alzheimer Center, a typical resident room or the outdoor therapeutic park.

By using your mouse or key.

The virtual tour is definitely worth a spin.

URL: www.heatherhill.com

VISUAL: A

WORTH VISITING: A

Test.com

Need some practice on taking test? Test.com is one Cleveland Web site that will help. Test.com is a venture-capital funded site that sells career planning, employment and other testing. Some tests are free, and for others there is a fee.

During February and in honor of Valentine's Day, there is a Love Quotient test for free. It's not as easy as you might think — unless you know your Shakespeare, movie trivia and history.

Some other sample tests are offered. The categories range from citizenship to typing. You answer by selecting a radio button or by filling in the blank. Clicking on another button scores the test and you receive the correct answers. You can even create your own test.

The site is light on graphics, but attractive.

URL: www.test.com

VISUAL: B+

WORTH VISITING: B+

Denny Carleton

Denny Carleton, musician and vocalist, has played with The Lost Souls, The Choir, Moses, The Pagans, Just Guise, and other rock bands popular in the 1960s, '70s and '80s in Greater Cleveland. He still plays music, but now in the coffee house and with a Christian theme.

He recently donated music, newspaper clippings and writings to the Western Reserve Historical Society. If you're into a trip down Memory Lane and the early days of rock music in Cleveland, read the history section of his Web site.

Women's Hormone Center

The Women's Hormone Center Beachwood offers answers by mail to medical concerns.

Other information files are line, and for further details, a book is available. Excerpts are quoted line. You can also turn to the page of links and find a dozen other sites on hormones and endocrinology.

The site is mostly text.

URL: www.hormonecenter.com

VISUAL: B-

WORTH VISITING: B-

Lake Erie Quality Wine Alliance

Fourteen Lake Erie wineries are listed on this Web site, and a tour ready for the taking.

Each of the fourteen has either full Web site or at least one information page. They are Buccia Vineyards, Chalet Debonne Vineyard, Conneaut Cellars Winery, Farnlands Winery, Klingshtrn Winery, Lonz Winery, Markko Vineyard, Mazza Vineyards, Mon Ami Restaurant and Historic Winery, Ol Firehouse Winery, Penn Shore Vineyards, Presque Isle Wine Cellars and Woodbury Vineyards.

Their locations range from the Lake Erie Islands to the shores of New York State. Some offer bed and breakfast inns and entertainment.

Chalet Debonne Vineyard, while a little short on photos, has a fine Web site among this group.

URL: www.rimestock.com/leqwa/index.html

VISUAL: B

WORTH VISITING: B

Goblin Bee Records

Goblin Bee Records is a cooperative jazz and blues record label whose performers are from Greater Cleveland. Eight titles have been

AT HOME

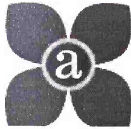
[back to LIFE](#)

Web Informant

- *Helps user evaluate quality of Web sites*
- *Identifies related sites of interest*
- *Can be downloaded free of charge*



Wednesday, March 18, 1998 - 11:04 AM ET



CBS

(CBS)

What It Is

Alexa 1.3, a browser aid which reveals information about the creator, or at least the host, of any given Web page, related sites, and more.

What It Does

Ever wonder who owns the page you're looking at? Or whether the information is any good?

While Alexa can't check facts for you, the software does pop up a slim menu bar which contains information about the sites a user visits, and acts as a kind of site authenticator and navigational aid.

Alexa puts two particularly powerful tools in the hands of its users: First, the company has collected and distilled a great deal of information about what sites people visit before and after they visit a given site. These "footpaths" of previous users often turn up lesser-known but high-quality and useful sites.

Alexa also has an archive feature. So, if one hits a page which is no longer connected to the Web – the infamous "404: File Not Found" error – there is a good chance Alexa has preserved a copy of the page in its memory banks.

The tool makes it easier to evaluate the quantity, freshness, and popularity of the sites you visit by noting the number of pages the site has, when they were last updated on average, and the relative amount of traffic that site receives.

In addition, Alexa features site ratings, currently from Yahoo! Internet Life and the Britannica Internet Guide. Privacy and content ratings are provided by online privacy advocate TRUSTe and content labeling organization RSACi.

What It Means

Alexa posits a future in which users know the source of the information they're looking at and have easy access to related content.

It also envisions a community of millions of Alexa users who, with their individual efforts, help to map and rate the Web for each other.

What It Costs

The software can be downloaded free of charge at Alexa's Web site, listed below.

[Smile For The Camera >>](#)



[Back to Spring Internet World](#)

Written by Sean Wolfe with graphic design by Dana Byerly.

RELATED LINKS

- [Alexa Internet](#)
- [TRUSTe member directory](#)
- [Recreational Software Advisory Council on the Internet \[RSACi\]](#)

[▲ back to TOP](#)

[[Home](#) | [News](#) | [Weather](#) | [Sports](#) | [Life](#) | [Local Guide](#) | [On Air](#) | [CBS.com](#)]
[[Feedback](#) | [Help](#)]

© 1998, CBS Worldwide Inc. & CBS Mass Media Corporation,
All Rights Reserved.

Please [click here](#) for more copyright information.



8C-0190

ONLY NEWSPAPER
THE ITEM
Seller: 8C

Circ- (R) 23,777 (R) 20,480

MARCH 6, 1998



Bacon's

3325-A

Searching and finding the information you really want on the Internet

Much of the e-mail I receive are questions dealing with how to search the Internet and obtain information. We are only in the fourth year of the public Internet, and we are literally drowning in information.

In previous articles, I attempted to provide some methods to find information that you, the reader, could use. In the past six months, we have seen a fourfold increase in search engines, and more than 100 million files available at FTP sites around the world. I have updated at www.sumter.sc.us the search engines, and made suggestions on which to use for

various searches. I recently found a site that helps find business applications, browser plug-ins, graphic images, sounds, fonts, screen savers, etc.

This site is located at <http://www.filez.com>, and provides other ways to find selected data by categories. This site is worth your visit. On the home page of this site, you will find the File of the Day. The particular day that I surfered the site, I found an image viewer that can be downloaded. In addition, it had a listing of the Previous Files of the Day; New Files encountered in the past two days, as well

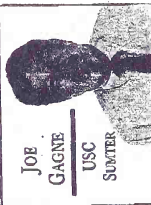
as a listing of the Top 20 Downloaded Applications, and the Top 20 Downloaded Games. This site also contains extensive category breakdowns such as New and Corporate Sites, Archives and Publications, Vendor Documentation, etc.

For instance, if you click the Vendor Documentation, you can obtain product information from Microsoft, Epson, and many other companies. The archives contain information about books and magazines, and even contain articles by subject matter. In addition, you can specify a file name or a description of a file, select the operating system (i.e., Windows 95), specify the number of sites you wish to find (5, 10, 20, 50, etc.) and click to search.

This site also contains a link to a powerful tool that you may wish to obtain that is free and usable with

either Netscape or Internet Explorer. This new tool is called [Filez](http://www.filez.com), and when installed on your system, creates a toolbar on the bottom of your browser. This toolbar will give you additional information when you are surfing about the current page you are on, and then gives appropriate references to pages that might be of interest to you. If you hit a site that reports a 404 unavailable site, I will seek outside information on the site's new location. In addition, it has on the toolbar access to an on-line encyclopedia, dictionary, thesaurus. When you download this software, be aware that the file is 1.6MB in size. It is an outstanding tool.

More on searching the Net next week. Questions/commands may be directed to joe-gagne@sumter.net. Good surfing to you.



PRODUCT REVIEWS

Web navigation utility

Alexa's free Web navigator deftly searches for Web sites

By Howard Millman

IF YOU'RE TIRED of sifting through the infotrash increasingly returned by standard Internet search engines, you may want to try Alexa Internet's free browser add-on, Alexa 1.4. This product can point you to Web sites with quality information and then suggest other sites that may offer similar or additional data. In my trials, Alexa made about an equal number of worthwhile and worthless recommendations, which still put me ahead of the game.

For example, I wanted to purchase a 35mm camera. I visited <http://www.allcamera.com>, a site that sells new and used cameras. Alexa dutifully suggested other photography sites to visit next. However, because Alexa did not know that

I wanted to purchase a camera, it suggested sites that generically related to photography but not camera sales. That's understandable because Alexa, and its primary competitors WiseWise and Firefly, could interpret my actions but not my purpose.

Alexa partially bases its conclusions about a site's content on its reading of the meta tags in a site's header. Consequently, its suggestions' accuracy relate to the honesty and competence of the site's owners.

Crawling the Web

Alexa's knowledge of a site's content results from Alexa Internet's ambitious program to crawl and catalog the Web. This mind-numbing operation offers a valuable side benefit to Alexa users. If a Web site or a page you want no longer exists, the tool



ALEXA'S TOOLBAR sits at the bottom of your browser. When clicked, it pops up suggestions on where to go next, as well as site statistics.

can deliver it up from its archive, providing it has crawled that site. When I tried to resurrect a site, I had to wait about 15 minutes as Alexa ground through its tape archives, but the product delivered. To date, Alexa's voracious spider has crawled more than 650,000 Web sites.

If Alexa's spider has not crawled a site, then that site's ratings and content analyses depend solely on users who have visited and evaluated the site. A site owner can request that Alexa include or exclude a site by filing a request at Alexa's Web site.

As well as navigating and making

content recommendations, Alexa reveals other site information. For example, a composite rating indicates how other visitors liked the site and who the site owner is, and provides statistics assessing its overall performance and the freshness of its content. You can add your opinion of the site by clicking on a button.

Alexa suffers from some technical limitations. For example, if it's used behind a corporate firewall, proxy settings must be reconfigured to let its data exchanger continue unimpeded.

Unobtrusive presence

Alexa limits its presence on your desktop to an unobtrusive toolbar. The toolbar also carries ads, which pay the freight, and free links to standard reference materials such as the *Encyclopedia Britannica* and the *Merriam-Webster* dictionary. A chat feature lets you conference online with Alexa-using colleagues.

I recommend that you use Alexa to supplement, not supplant, a tra-

THE BOTTOM LINE



Alexa 1.4

This simple Web navigator sits at the bottom of your browser. It seeks to improve on traditional search engines by combining a collaborative filtering technique and grading system with a search engine.

- **Pros:** Free; easy to use.
- **Cons:** Suggestions varied from a perfect match to totally irrelevant; tends to overlook smaller sites.

▼ **Alexa Internet Inc., San Francisco;**
(800) 882-5392 (toll-free), (415)
561-4900; fax: (415) 561-6795;
<http://www.alexa.com>

- **Price:** Free.
- **Platforms:** Windows 95, Windows NT 4.0.

ditional search engine. But as Alexa divines your research goals, it will minimize the need to sift through the vast amount of extraneous data.

Howard Millman (hmillman@ibm.net) operates the Data System Service Group, in Croton, N.Y.

Hey, Nobody Said You Had To Share.
Get Your Very Own InfoWorld Subscription.

www.iwsubscribe.com

www.iwsubscribe.com

Make sure you get exactly what you want out of life. InfoWorld has weekly issues with your name on it. Just apply online. It's fast and the subscription is free if you qualify. Visit the InfoWorld Web site today.

INFO
WORLD

WORKING WOMAN

May 1998

Your Cyberspace Trail Guide

A software start-up profits by archiving the Web's most popular pathways.

IF YOU'VE EVER FELT LOST ON the World Wide Web, a pioneering Internet scientist has designed a new software product for you. Just as hikers benefit from old trekking paths, Web surfers using Alexa Internet can follow the trails of thousands of browsers who preceded them.

Alexa is a navigator that works in tandem with Netscape or Internet Explorer. (Paid for by advertisers, it is free of charge to users.) "I think of it as a co-pilot for wandering around cyberspace," says Jerry Michalski, a technology analyst at the newsletter *Release 1.0*, who predicts it will be a hit.

Insiders call Alexa's creator, Brewster Kahle, 37, a brilliant visionary. "Everything he touches turns to gold," says Joella Paquette, marketing development manager at Digital Equipment Corp.'s Network Systems Laboratory. Designing supercomputers shortly after graduating from Massachusetts Institute of Technology, Kahle then created some of the earliest Internet publishing tools and production services that enabled the government and private firms to get on the Net. America Online bought him out in 1995.

Alexa, his latest brainstorm, can:

- Locate extinguished links whenever you hit

a dead-end on the Web, retrieving and displaying Web pages that no longer exist;

- Suggest where to go next on the Web, based on the records of previous visitors;

- Rank the site you're visiting and display information about its owner and contents.

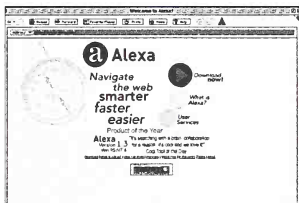
Alexa also donates everything it finds to a not-for-profit information warehouse Kahle recently founded. The Seattle-based Internet Archive is a digital library that's amassing every publicly accessible Web page, newsgroup posting, and file that's ever flashed into existence. Buried in this mountain

of data are the pathways trod by millions of Internet users, which provide valuable clues to how people use the Net, part of Kahle's historical mission.

Because the life span of the average Web page is just 75 days, our "digital heritage" is disappearing fast, Kahle explains. "The archive will be seen a century from now as an unprecedented collection of the human voice." It's already eight terabytes big—that's one trillion bytes—and growing.

Download Alexa at www.alexa.com. For now, you must run Windows 95 or Windows NT 4.0. A Macintosh-compatible version is in the works.

—PETER ZHEUTLIN



Technology | Tom Watson

Alternative doorways to the Internet are popping up in the spirit of free-flowing information.

THERE were only a few snickers last week at a new-media conference in Chicago when someone asked a group of executives from Web-search companies which of them would be the first to buy a major television network. It was a sign of just how quickly the likes of Yahoo, Lycos, Excite and Infoseek have grown, how grand their mass media dreams have become and how much their original vision of the World Wide Web has changed.

No longer content to direct a user to the Web sites most relevant to the individual's interests and tastes, the former search services are intent on creating "portals" to the Internet, complete with dial-in service, free E-mail and personal Web pages, paging and messaging, customized news, local weather and stock tracking. The aim: get the suckers under the tent and keep them there. That's a radical departure from the original business model of catching the customer's eye on the way to other sites.

The original search model was deeply steeped in the Internet ethos — the idea that information and traffic flows freely from site to site and value is added by serving that urge, not resisting it. It is an ethos that favors entropy over organization, an open range of Web publishing compared with the stockyard chutes of the portal sites. And it is an ethos that for many in the Web business still endures — despite the portal fad.

"There are between 5 and 10 million content areas on the Web, and in many ways the Internet is more like print publishing," observed Brewster Kahle.

Mr. Kahle is an Internet pioneer whose 1989 invention, the Wide Area Information Server — or WAIS — was a pre-Web system for searching distant data bases on the Internet. He later sold his on-line and software publishing company, WAIS Inc., to America Online.

The main portals are following a model that mimics the control and distribution of cable television networks. But Mr. Kahle thinks that is the wrong model. "There are 16,000 journal publishers in print, real diversity. And everyone's experience is different. Do we need a TV Guide? I think we need something a little more sophisticated."

Certainly the current search services are far from exhaustive. A study released last month by the NEC Re-

search Institute of Princeton, N.J., indicated that even the most thorough service, Hotbot (www.hotbot.com), has indexed only 34 percent of the Web's estimated 320 million pages.

But Mr. Kahle is hardly a disinterested observer. He is the president and co-founder with Bruce Gilliat of Alexa Internet, a two-year-old start-up company based in San Francisco. Their product is Alexa, part Web browser, part navigation service. Users download the software from www.alexa.com free, after which Alexa manifests itself as a thin toolbar under the regular Web browser — whether the PC is using the Netscape Navigator or Internet Explorer from Microsoft.

Alexa offers quick access to information on each site visited (who owns it, how much traffic it gets, and how Alexa users rated its content), provides links to other similar sites and, lately, tiny advertising messages keyed to the user's browsing selections. About 350,000 copies of the program have been downloaded and there are 100,000 regular users, according to Mr. Kahle.

Alexa is a geeky end-run around the sleek mass media dreams of the search engine companies. While Excite, Yahoo, Lycos and Infoseek are adding as many features as possible to keep users on their sites for as long as possible, Alexa encourages wide and frequent grazing by recommending sites wherever the user's interests may lead, based on the person's past preferences — and based on the preferences of other users who have frequented the same sites.

Mr. Kahle calls this approach "contextual navigation." The more users in Alexa's data base, the better the similar-preferences software works — and the more precisely that Alexa can tailor its ads to individual users. In other words, the more that users surf outside the main portals, the better Alexa's revenue stream.

Alexa is named for the library of Alexandria, the ill-fated attempt of the ancient Greeks to amass all of the literate world's printed knowledge. And in keeping with this ideal, Alexa brings the emphasis in Web navigation back to content — not just a reader's digest of the Web.

And yet, in choosing to name his venture after an ambitious idea that ultimately fell short, Mr. Kahle is im-



Stuart Goldenberg

PLICITLY conceding that the sheer size, growth, and second-to-second mutability of the Internet makes it almost impossible to amass the collected works of the Web. The reason the current search services consistently rank among the most popular Web sites is that many people presumably do want some winnowing.

But the mass-market model need not be the only portal approach. "We believe that people are more in need of filters than catch-alls," said Rufus Griscom, co-founder of the artily erotic on-line magazine Nerve (www.nervemag.com). That is why Nerve, which bills itself as "literate smut," has created its own, more narrow portal: a directory of sexually oriented Web sites.

And Nerve is not alone in providing a narrower doorway to the Web. Alternative portals are everywhere, including Razorfish's "Disinformation" search engine (www.disinfo.com) that provides links to various subculture sites, and "John Skilton's Baseball Links" (www.baseball-links.com), perhaps the most complete guide to baseball on the Web.

These alternate portals are evidence of the Internet ethos that refuses to conform to a mass media structure. In their race to emulate mainstream media giants like Time Warner, CBS and Disney, the search engines may be forgetting the very phenomenon that brought them into being: the Internet is a medium of creators as much as it is a medium for consumers.

Tom Watson is editor and co-founder of @NY (www.atnewyork.com), an information service that focuses on New York's interactive industry.

